On stability of difference schemes. Central schemes for hyperbolic conservation laws with source terms

M. Mond [*]   V. S. Borisov [†]

September 20, 2009

Abstract

The stability of difference schemes for, in general, hyperbolic systems of conservation laws with source terms are studied. The basic approach is to investigate the stability of a non-linear scheme in terms of its corresponding scheme in variations. Such an approach leads to application of the stability theory for linear equation systems to establish stability of the corresponding non-linear scheme. It is established the notion that a non-linear scheme is stable if and only if the corresponding scheme in variations is stable.

A new modification of the central Lax-Friedrichs (LxF) scheme is developed to be of the second order accuracy. A monotone piecewise cubic interpolation is used in the central schemes to give an accurate approximation for the model in question. The stability of the modified scheme are investigated. Some versions of the modified scheme are tested on several conservation laws, and the scheme is found to be accurate and robust.

As applied to hyperbolic conservation laws with, in general, stiff source terms, it is constructed a second order nonstaggered central scheme based on operator-splitting techniques.

1   Introduction

We are mainly concerned with the stability of difference schemes for hyperbolic systems of conservation laws with source terms. Such systems are used to describe many physical problems of great practical importance in magneto-hydrodynamics, kinetic theory of rarefied gases, linear and nonlinear waves, viscoelasticity, multi-phase flows and phase transitions, shallow waters, etc. (see,

---

[*]The Pearlstone Center for Aeronautical Engineering Studies, Department of Mechanical Engineering, Ben-Gurion University of the Negev, Beer-Sheva, Israel. E-mail: mond@bgu.ac.il

[†]E-mail: viatslav@bgu.ac.il

1

e.g., [6], [10], [17], [24], [30], [32], [34], [37], [41], [42]). We will consider a system of hyperbolic conservation laws written as follows (e.g., [17], [32])

$$\frac{\partial \mathbf{u}}{\partial t} + \sum_{j=1}^{N} \frac{\partial}{\partial x_j} \mathbf{f}_j(\mathbf{u}) = \frac{1}{\tau} \mathbf{q}(\mathbf{u}), \ 0 < t \le T_{\max}, \ \mathbf{u}(\mathbf{x}, t)|_{t=0} = \mathbf{u}^0(\mathbf{x}), \quad (1)$$

where $\mathbf{x} \equiv \{x_1, x_2, \ldots, x_N\}^T \in \mathbb{R}^N$, $\mathbf{u} = \{u_1, u_2, \ldots, u_M\}^T$ is a vector-valued function from $\mathbb{R}^N \times [0, +\infty)$ into an open subset $\Omega_{\mathbf{u}} \subset \mathbb{R}^M$, $\mathbf{f}_j(\mathbf{u}) = \{f_{1j}(\mathbf{u})$, $f_{2j}(\mathbf{u}), \ldots, f_{Mj}(\mathbf{u})\}^T$ is a smooth function (flux-function) from $\Omega_{\mathbf{u}}$ into $\mathbb{R}^M$, $\mathbf{q}(\mathbf{u}) = \{q_1(\mathbf{u}), q_2(\mathbf{u}), \ldots, q_M(\mathbf{u})\}^T$ denotes the source term, $\tau > 0$ denotes the stiffness parameter, $\mathbf{u}^0(\mathbf{x})$ is of compact support. We will assume that $\tau = const$ without loss of generality. In what follows $\|\mathbf{M}\|_p$ denotes the matrix norm of a matrix $\mathbf{M}$ induced by the vector norm $\|\mathbf{v}\|_p = (\sum_i |v_i|^p)^{1/p}$, and $\|\mathbf{M}\|$ denotes the matrix norm induced by a prescribed vector norm. $\mathbb{R}$ denotes the field of real numbers.

For studying stability and monotonicity of non-linear schemes, the well known notion of total variation diminishing (TVD, see, e.g., [17], [32]) turns out to be an useful tool. Actually, the following property

$$\|\mathcal{N}(\mathbf{v} + \delta\mathbf{v}) - \mathcal{N}(\mathbf{v})\| \le (1 + \alpha \Delta t) \|\delta\mathbf{v}\| \qquad (2)$$

is sufficient for stability of a two-step method [32], however it is, in general, difficult to obtain. Here Δt denotes the time increment, α is a constant independent of Δt as $\Delta t \to 0$, $\mathbf{v}$ and $\delta\mathbf{v}$ are any two grid functions ($\delta\mathbf{v}$ will often be referred to as the variation of the grid function $\mathbf{v}$), $\mathcal{N}$ denotes the scheme operator. At the same time, the stability of linearized version of the non-linear scheme is generally not sufficient to prove convergence [20], [32]. Instead, the TV-stability adopted in [20] (see also [32, s. 8.3.5]) makes it possible to prove convergence (to say, TV-convergence) of non-linear scalar schemes with ease. However, the TVD property is a purely scalar notion that cannot, in general, be extended for non-linear systems of equations, as the true solution itself is usually not TVD [17], [32]. Moreover, one can see in [8, pp. 1578-1581] that a TVD scheme can be non-convergent in, at least, L_∞, in spite that the scheme is TV-stable. Such a phenomenon is, in all likelihood, caused by the fact that TV is not a norm, but a semi-norm.

Nowadays, there exists a few methods for stability analysis of some classes of nonlinear difference schemes approximating systems of PDEs (see, e.g., [14], [16], [32], [35], [38], [48] and references therein). It is noted in [16] that the problem of stability analysis is still one of the most burning problems, because of the absence of its complete solution. In particular, as noted in [14] in this connection, the vast majority of difference schemes, currently in use, have still not been analyzed. LeVeque [32] noted as well that, in general, no numerical method for non-linear systems of equations has been proven to be stable. There is not even a proof that the first-order Godunov method converges on general systems of non-linear conservation laws [32, p. 340]. Thus, a different approach

to testing scheme stability must be adopted to prove convergence of non-linear schemes for systems of PDEs. The notion of scheme in variations (or variational scheme [8], [9]) has, in all likelihood, much potential to be an effective tool for studying stability of nonlinear schemes. Such an approach goes back to the one suggested by Lyapunov (1892), namely, to investigate stability by the first approximation. This idea has long been exploited for investigation of the stability of motion [15]. An approach to investigate non-linear difference schemes for monotonicity in terms of corresponding variational schemes was suggested in [8], [9]. The advantage of such an approach is that the variational scheme will always be linear and, hence, enables the investigation of the monotonicity for nonlinear operators using linear patterns. It is proven for the case of explicit schemes that the monotonicity of a variational scheme will guarantee that its original scheme will also be monotone [8]. We establish the notion that the stability of a scheme in variations is necessary and sufficient for the stability of its original scheme (see Section 2, Theorem 4).

An extensive literature is devoted to central schemes, since these schemes are attractive for various reasons: no Riemann solvers, characteristic decompositions, complicated flux splittings, etc., must be involved in construction of a central scheme (see, e.g., [5], [30], [31], [32], [41], [43] and references therein), and hence such schemes can be implemented as a black-box solvers for general systems of conservation laws [30]. Let us, however, note that the numerical domain of dependence [32, p. 69] for a central scheme approximating, e.g., a scalar transport equation coincides with the numerical domain of dependence for a standard explicit scheme approximating diffusion equations [32, p. 67]. Such a property is inherent to central schemes in contrast to, e.g., the first-order upwind schemes [32, p. 73]. Hence, central schemes do not satisfy the long known principle (e.g., [2, p. 304]) that derivatives must be correctly treated using type-dependent differences, and hence there is a risk for every central scheme to exhibit spurious solutions. The results of simulations in [39] can be seen as an illustration of the last assertion. Notice, all versions of the, so called, Nessyahu-Tadmor (NT) central scheme, in spite of sufficiently small CFL (Courant-Friedrichs-Lewy [32]) number ($Cr = 0.475$), exhibit spurious oscillations in contrast to the second-order upwind scheme ($Cr = 0.95$). The first order, $O(\Delta t + \Delta x)$, LxF scheme exhibits the excessive numerical viscosity. Thus, the central scheme should be chosen with great care to reflect the true solution and to avoid significant but spurious peculiarities in numerical solutions.

Let us note that LxF scheme – the forerunner for central schemes [5], [30] – does not produce spurious oscillations. While, from the pioneering works of Nessyahu and Tadmor [39] and on, the higher order versions of LxF scheme can produce spurious oscillations. The reason has to do with a negative numerical viscosity introduced to obtain a higher order accurate scheme (for more details, see Section 4). Let us note that there is a possibility to increase the scheme's order of accuracy, up to $O((\Delta t)^2 + (\Delta x)^2)$, by introducing an additional non-negative numerical viscosity into the scheme. Such an approach is similar to the vanishing viscosity method [17], [32], and hence possesses its advantages, yet

3

it appears to be free of the disadvantages of this method, since the additional viscosity term is not artificial. With this approach, the second order scheme is developed in Section 4, where sufficient conditions for stability of the scheme are found. The scheme is tested on several conservation laws in Section 5.

A stable numerical scheme may yield spurious results when applied to a stiff hyperbolic system with relaxation (see, e.g., [1], [4], [6], [10], [11], [24], [44], [45]). Specifically, spurious numerical solution phenomena may occur when underresolved numerical schemes (i.e., insufficient spatial and temporal resolution) are used (e.g., [1], [24], [26], [37]). However, during a computation, the stiffness parameter may be very small, and, hence, to resolve the small stiffness parameter, we need a huge number of time and spatial increments, making the computation impractical. Hence, we are interested to solve the system, (1), with underresolved numerical schemes. It is significant that for relaxation systems a numerical scheme must possess a discrete analogy to the continuous asymptotic limit, because any scheme violating the correct asymptotic limit leads to spurious or poor solutions (see, e.g., [10], [24], [25], [37], [41]). Most methods for solving such systems can be described as operator splitting ones, [11], or methods of fractional steps, [6]. After operator splitting, one solves the advection homogeneous system, and then the ordinary differential equations associated with the source terms. As reported in [18], this approach is well suited for the stiff systems. We are mainly concerned with such an approach in Section 4.2.

2 Stability of difference schemes

Let us consider the following non-linear explicit scheme arising, e.g., in numerical analysis of nonlinear PDE systems:

$$\mathbf{v}_i^{n+1} = \mathbf{H}_i^n(\mathbf{v}_1^n, \mathbf{v}_2^n, \ldots, \mathbf{v}_I^n), \ \mathbf{H}_i^n : \Omega_n \subseteq \mathbb{R}^N \to \mathbb{R}^{N_0}, \ i \in \omega_1, \ n, n+1 \in \omega_2, \ (3)$$

where $\mathbf{v}_i^n \in \mathbb{R}^{N_0}$ denotes a vector-valued grid function, $N = N_0 I$, $i \in \omega_1$ denotes a node of the grid $\omega_1 \equiv \{1, 2, \ldots, I\}$, $n \in \omega_2$ denotes a node (time level) of the grid $\omega_2 \equiv \{0, 1, \ldots, M\}$, $\mathbf{H}_i^n = \left\{ H_{i,1}^n, \ H_{i,2}^n, \ \ldots, \ H_{i,N_0}^n \right\}^T$ is a vector-valued function with the domain and range belonging to $\mathbb{R}^N$ and $\mathbb{R}^{N_0}$, respectively. Notice, $\mathbf{H}_i^n$ depends also on scheme parameters (e.g., space and time increments), however, this dependence is usually not included in the notation. We will assume that n in (3) denotes the time level, $t_n \ (= n\Delta t)$. Thus, the time increment will be represented by $\Delta t = t_{\max}/M = const$, where $t_{\max}$ denotes some finite time over which we wish to compute. If we introduce the additional notation

$$\mathbf{v}^n = \left\{ (\mathbf{v}_1^n)^T, (\mathbf{v}_2^n)^T, \ldots, (\mathbf{v}_I^n)^T \right\}^T, \ \mathbf{H}^n = \left\{ (\mathbf{H}_1^n)^T, (\mathbf{H}_2^n)^T, \ldots, (\mathbf{H}_I^n)^T \right\}^T, \ (4)$$

then the scheme (3) can be written in the form

$$\mathbf{v}^{n+1} = \mathbf{H}^n(\mathbf{v}^n), \quad \mathbf{H}^n : \Omega_n \subseteq \mathbb{R}^N \to \mathbb{R}^N, \quad n, n+1 \in \omega_2 \equiv \{0, 1, \ldots, M\}. \ (5)$$

As usual (e.g., [40, p. 62]), for mappings $\mathbf{f} : \Omega_f \subseteq \mathbb{R}^N \to \mathbb{R}^N$ and $\mathbf{g} : \Omega_g \subseteq \mathbb{R}^N \to \mathbb{R}^N$, the composite mapping $\mathbf{h} = \mathbf{g} \circ \mathbf{f}$ is defined by $\mathbf{h}(\mathbf{v}) = \mathbf{g}(\mathbf{f}(\mathbf{v}))$ for

all $\mathbf{v} \in \Omega_h = \{\mathbf{v} \in \Omega_f \mid \mathbf{f}(\mathbf{v}) \in \Omega_g\}$. Using the composite mapping approach, we rewrite Scheme (5) to read

$$\mathbf{y} = \mathbf{F}(\mathbf{x}), \quad \mathbf{F} : \Omega_F \subseteq \mathbb{R}^N \to \mathbb{R}^N, \tag{6}$$

where the following notation is used: $\mathbf{x} = \mathbf{v}^0$, $\mathbf{y} = \mathbf{v}^M$, $\mathbf{F} = \mathbf{H}^{M-1} \circ \mathbf{H}^{M-2} \circ \ldots \circ \mathbf{H}^0$, $\Omega_F = \{\mathbf{v}^0 \in \Omega_0 \mid \mathbf{v}^1 = \mathbf{H}^0(\mathbf{v}^0) \in \Omega_1 \mid \ldots \mid \mathbf{v}^{M-1} = \mathbf{H}^{M-2}(\mathbf{v}^{M-2}) \in \Omega_{M-1}\}$. Let the scheme parameters (including time increments) be represented by a vector $\mathbf{s}$ belonging to some normed space with the norm $|\mathbf{s}|$.

Since differentiability of $\mathbf{H}^n$ as well as $\mathbf{F}$ will be a key element in the following, let us note that the composite mapping $\mathbf{F}$ will be Fréchet-differentiable [40, item 3.1.5] if all of the maps, $\mathbf{H}^n$, are Fréchet-differentiable [40, item 3.1.7]. However, if all of the maps are Fréchet-differentiable, but one that Gateaux-differentiable [40, item 3.1.1], then the composite mapping $\mathbf{F}$ has a Gateaux-derivative [40, item 3.1.7]. Notice, if there exist at least two maps having Gateaux-derivatives, then $\mathbf{F}$ need not be differentiable [40, E 3.l-7].

Scheme (6) is said to be stable (see, e.g., [14], [17], [32], [47], [48], [49]) if there exist positive s_0, $C = const$ such that for all $\mathbf{x}$, $\mathbf{x}_* \in \Omega_F$ the following inequality is valid

$$\|\mathbf{F}(\mathbf{x}_*) - \mathbf{F}(\mathbf{x})\| \leq C \|\mathbf{x}_* - \mathbf{x}\|, \quad \forall \, \mathbf{s} : \ |\mathbf{s}| \leq s_0. \tag{7}$$

Thus, Scheme (6) will be stable *iff* (if and only if) the function $\mathbf{F}$ will be Lipschitz for a constant C.

To be more specific, let us consider the "slit plane" [22] in polar coordinates (r, θ)

$$\Omega_F = \{(r, \theta) \mid 0 < r < \infty, \ -\pi < \theta < \pi\} \subset \mathbb{R}^2, \tag{8}$$

and the function $\mathbf{F} = \{F_1(r, \theta), F_2(r, \theta)\}^T$ such that [22]

$$F_1 = r, \ F_2 = \theta \big/ 2. \tag{9}$$

If we take $(r, \theta)_* = (r_0, -\pi + \varepsilon)$, $(r, \theta) = (r_0, \pi - \varepsilon)$, and $r_0 = const$, then obviously the mapping (9) is not Lipschitz, since C in (7) tends to infinity as $\varepsilon \to 0$. Therefore, we have to conclude, in view of the above definition, that Scheme (9) is not stable, even though the function $\mathbf{F}$, (9), is locally Lipschitz for $C = 1$, and, further, $\mathbf{F}$ is the non-stretching mapping of the "slit plane" (8) into the right semi-plane. Hence, the preceding definition of stability needs minor changes.

A set $\Omega \subseteq \mathbb{R}^N$ is said to be path-connected if every two points $\mathbf{x}$, $\mathbf{x}_* \in \Omega$ can be joined by a continuous curve ($\gamma : [0, 1] \subset \mathbb{R} \to \Omega$, [22], [29, p. 113]) of finite length, $L(\gamma)$. The intrinsic metric [22] Λ_Ω in a path-connected set Ω is defined as

$$\Lambda_\Omega(\mathbf{x}, \mathbf{x}_*) = \inf_{\gamma \subset \Omega} L(\gamma), \quad \gamma : \ \mathbf{x} = \gamma(0), \ \mathbf{x}_* = \gamma(1), \ L(\gamma) < \infty. \tag{10}$$

An open ball (of radius r) about $\mathbf{x} \in \mathbb{R}^N$ is denoted by $B(\mathbf{x}, r)$ (or just $B_\mathbf{x}$).

Definition 1 *Let Ω_F in (6) be path-connected. Scheme (6) is said to be stable if there exist positive s_0, $C = const$ such that the following inequality holds*

$$\|\mathbf{F}(\mathbf{x}_*) - \mathbf{F}(\mathbf{x})\| \leq C\Lambda_{\Omega_F}(\mathbf{x}, \mathbf{x}_*), \quad \forall\, \mathbf{x}, \mathbf{x}_* \in \Omega_F, \quad \forall\, \mathbf{s}: \ |\mathbf{s}| \leq s_0. \quad (11)$$

Notice, Scheme (9) is stable, since Inequality (11) holds for $C = 1$.

Lemma 2 *Let the path-connected Ω_F of (6) be open in $\mathbb{R}^N$. Scheme (6) will be stable in terms of Definition 1* iff *$\mathbf{F}$ in (6) will be locally Lipschitz for a common constant C, for all scheme parameters $\mathbf{s}$ such that $|\mathbf{s}| \leq s_0$.*

Proof. Suppose Scheme (6) is stable, i.e. (11) is valid. Choose any point $\mathbf{x} \in \Omega_F$. Since Ω_F is open, there exists a radius r such that $B(\mathbf{x}, r) \subset \Omega_F$. Choose any point $\mathbf{x}_* \in B(\mathbf{x}, r)$, and let γ_* be the straight line segment joining the points $\mathbf{x}, \mathbf{x}_* \in B(\mathbf{x}, r)$. In view of (11), $\mathbf{F}$ in (6) will be locally Lipschitz for a common constant C, for all $\mathbf{s}: \ |\mathbf{s}| \leq s_0$, since $\Lambda_{\Omega_F}(\mathbf{x}, \mathbf{x}_*) = L(\gamma_*) = \|\mathbf{x}_* - \mathbf{x}\|$.

Conversely, suppose that $\mathbf{F}$ in (6) is locally Lipschitz for a common constant C, for all $\mathbf{s}: \ |\mathbf{s}| \leq s_0$. Let some points $\mathbf{x}, \mathbf{x}_* \in \Omega_F$ be joined by a continuous curve γ. In view of (10), the curve γ can be taken such that $L(\gamma) \leq \Lambda_{\Omega_F}(\mathbf{x}, \mathbf{x}_*) + \varepsilon$ for an arbitrary $\varepsilon > 0$. Given any point $\mathbf{z} \in \gamma$, there is a ball $B_\mathbf{z} \subset \Omega_F$. The balls $\{B_\mathbf{z}\}$ form an open cover of γ. Since the mapping $\gamma : [0, 1] \subset \mathbb{R} \to \mathbb{R}^N$ is continuous, the curve γ is compact [29, p. 94]. Hence, by the compactness of γ, $\{B_\mathbf{z}\}$ has a finite subcover consisting of balls $B_\mathbf{x} = B_{\mathbf{z}_1}, B_{\mathbf{z}_2}, \ldots, B_{\mathbf{z}_K} = B_{\mathbf{x}_*}$. Since $\mathbf{F}$ is locally Lipschitz, we find

$$\|\mathbf{F}(\mathbf{z}_{k+1}) - \mathbf{F}(\mathbf{z}_k)\| \leq C\|\mathbf{z}_{k+1} - \mathbf{z}_k\|, \quad k = 1, 2, \ldots K-1, \ \forall\, \mathbf{s}: \ |\mathbf{s}| \leq s_0. \ (12)$$

Then, by virtue of (12), we find

$$\|\mathbf{F}(\mathbf{x}_*) - \mathbf{F}(\mathbf{x})\| = \left\|\sum_k [\mathbf{F}(\mathbf{z}_{k+1}) - \mathbf{F}(\mathbf{z}_k)]\right\| \leq C\sum_k \|\mathbf{z}_{k+1} - \mathbf{z}_k\| \leq$$

$$CL(\gamma) \leq C\Lambda_{\Omega_F}(\mathbf{x}, \mathbf{x}_*) + \varepsilon C, \quad \forall\, \mathbf{s}: \ |\mathbf{s}| \leq s_0. \quad (13)$$

By letting $\varepsilon \to 0$, we find that (11) holds. ∎

Let us find the necessary and sufficient conditions for the stability of Scheme (6). Let $W^{1,\infty}(\Omega_F)$ denote the Sobolev space, and let $\mathbf{F} \equiv \{F_1, F_2, \ldots, F_N\}^T$ in (6). Then, F_i, $i = 1, 2, \ldots, N$, (and, hence, $\mathbf{F}$) is locally Lipschitz (in the sense of having representatives) *iff* $F_i \in W^{1,\infty}(\Omega_F)$ (see, e.g., [22, Theorem 4.1]). Let ∇F_i denote the distributional gradient of F_i, and let $\delta\mathbf{F}$, $\delta\mathbf{x} \in \mathbb{R}^N$ denote variations. The following equality

$$\delta\mathbf{F} = \mathbf{F}' \cdot \delta\mathbf{x}, \quad \mathbf{F}' = \{\nabla F_1, \nabla F_2, \ldots, \nabla F_N\}^T, \quad (14)$$

will be viewed as the scheme in variations for (6).

Lemma 3 *Linear Scheme (14) will be stable* iff *there exist positive s_0, $C = const$ such that*

$$\|\mathbf{F}'\| \leq C = const, \quad \forall\, \mathbf{x} \in \Omega_F, \quad \forall\, \mathbf{s}: \ |\mathbf{s}| \leq s_0. \quad (15)$$

Proof. The sufficiency is obvious. Actually, by virtue of (15), we find that $\|\delta \mathbf{F}\| = \|\mathbf{F}' \cdot \delta \mathbf{x}\| \leq \|\mathbf{F}'\| \, \|\delta \mathbf{x}\| \leq C \, \|\delta \mathbf{x}\|$, i.e.

$$\|\delta \mathbf{F}\| \leq C \, \|\delta \mathbf{x}\|, \quad \forall \, \mathbf{x} \in \Omega_F, \quad \forall \, \mathbf{s} : \, |\mathbf{s}| \leq s_0. \tag{16}$$

Conversely, suppose that (16) is valid. Then, in view of [29, Theorem 2, p. 224], we write

$$\|\mathbf{F}'\| = \sup_{\|\delta \mathbf{x}\| \neq 0} \frac{\|\mathbf{F}' \cdot \delta \mathbf{x}\|}{\|\delta \mathbf{x}\|} = \sup_{\|\delta \mathbf{x}\| \neq 0} \frac{\|\delta \mathbf{F}\|}{\|\delta \mathbf{x}\|} \leq \sup_{\|\delta \mathbf{x}\| \neq 0} \frac{C \, \|\delta \mathbf{x}\|}{\|\delta \mathbf{x}\|} = C. \tag{17}$$

Hence, (15) holds, in view of (17) ∎

Theorem 4 *Consider Scheme (6). Let the path-connected Ω_F be open, $\mathbf{F} \equiv \{F_1 , F_2, \ldots, F_N\}^T$ be bounded, and let $\|\mathbf{F}'\| \equiv \|\mathbf{f}_F\|$, $\mathbf{f}_F \equiv \{\|\nabla F_1\|_\infty , \|\nabla F_2\|_\infty , \ldots, \|\nabla F_N\|_\infty\}^T$, ∇F_i, $i = 1, 2, \ldots, N$, denote the distributional gradient of F_i. Then, Scheme (6) will be stable* iff *its scheme in variations, (14), will be stable.*

Proof. The proof is trivial. Actually, Scheme (6) is stable $\Longleftrightarrow$ $\mathbf{F}$ is locally Lipschitz for a common constant C (Lemma 2) $\Longleftrightarrow$ $F_i \in W^{1,\infty}(\Omega_F)$ (see [22, Theorem 4.1]) $\Longleftrightarrow$ (15) holds $\Longleftrightarrow$ Scheme in variations, (14), is stable (Lemma 3). ∎

Notice, if $\mathbf{F}$ in (6) is Gateaux-differentiable, then ∇F_i (see Theorem 4) denotes the classical gradient of F_i, and, hence, it may be taken that $\mathbf{f}_F = \{\|\nabla F_1\| , \|\nabla F_2\|, \ldots, \|\nabla F_N\|\}^T$, see also [7, Theorem 3].

3 Monotone C^1 piecewise cubics in construction of central schemes

In this section we consider some theoretical aspects for high-order interpolation and employment of monotone C^1 piecewise cubics (e.g., [12], [28]) in construction of monotone central schemes. We will consider explicit schemes on a uniform grid with time step Δt and spatial mesh size Δx, as applied to the following hyperbolic 1-D equation

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial}{\partial x} \mathbf{f}(\mathbf{u}) = 0, \; t_n < t \leq t_{n+1} \equiv t_n + \Delta t, \quad \mathbf{u}(x, t_n) = \mathbf{u}^n(x), \tag{18}$$

Using the central differencing, we write

$$\left. \frac{\partial \mathbf{u}}{\partial t} \right|_{t=t_{n+0.25}, \; x=x_{i+0.5}} = \frac{\mathbf{u}_{i+0.5}^{n+0.5} - \mathbf{u}_{i+0.5}^n}{0.5 \Delta t} + O\left((\Delta t)^2\right), \tag{19}$$

$$\left. \frac{\partial \mathbf{f}}{\partial x} \right|_{t=t_{n+0.25}, \; x=x_{i+0.5}} = \frac{\mathbf{f}_{i+1}^{n+0.25} - \mathbf{f}_i^{n+0.25}}{\Delta x} + O\left((\Delta x)^2\right). \tag{20}$$

By virtue of (19)-(20) we approximate (18) on the cell $[x_i, x_{i+1}] \times [t_n, t_{n+0.5}]$ by the following difference equation

$$\mathbf{v}_{i+0.5}^{n+0.5} = \mathbf{v}_{i+0.5}^n - \frac{\Delta t}{2\Delta x} \left(\mathbf{g}_{i+1}^{n+0.25} - \mathbf{g}_i^{n+0.25} \right). \tag{21}$$

As usual, the mathematical treatment for the second step (i.e., on the cell $[x_{i-0.5}, x_{i+0.5}] \times [t_{n+0.5}, t_{n+1}]$) of a staggered scheme will, in general, not be included in the text, because it is quite similar to the one for the first step.

Considering that (21) approximates (18) with the accuracy $O((\Delta x)^2 + (\Delta t)^2)$, the next problem is to approximate $\mathbf{v}_{i+0.5}^n$ and $\mathbf{g}_i^{n+0.25}$ in such a way as to retain the accuracy of the approximation. For instance, the following approximations

$$\mathbf{v}_{i+0.5}^n = 0.5 \left(\mathbf{v}_i^n + \mathbf{v}_{i+1}^n \right) + O\left((\Delta x)^2 \right), \quad \mathbf{g}_i^{n+0.25} = \mathbf{f}\left(\mathbf{v}_i^n \right) + O\left(\Delta t \right), \tag{22}$$

leads to the staggered form of the famed LxF scheme that is of the first-order approximation (see, e.g., [17, p. 170]). One way to obtain a higher-order scheme is to use a higher order interpolation. At the same time it is required of the interpolant to be monotonicity preserving. Notice, the classic cubic spline does not possess such a property (see Figure 1a). Let us consider the problem of high-order interpolation of $\mathbf{v}_{i+0.5}^n$ in (21) with closer inspection

Let $\mathbf{p} = \mathbf{p}(x) \equiv \left\{ p^1(x), \dots, p^k(x), \dots, p^m(x) \right\}^T$ be a component-wise monotone C^1 piecewise cubic interpolant (e.g., [12], [28]), and let

$$\mathbf{p}_i = \mathbf{p}(x_i), \quad \mathbf{p}_i' = \mathbf{p}'(x_i), \quad \Delta \mathbf{p}_i = \mathbf{p}_{i+1} - \mathbf{p}_i,$$

$$\mathbf{p}_i' = \mathbb{A}_i \cdot \frac{\Delta \mathbf{p}_i}{\Delta x}, \quad \mathbf{p}_{i+1}' = \mathbb{B}_i \cdot \frac{\Delta \mathbf{p}_i}{\Delta x}, \tag{23}$$

where $\mathbf{p}_i'$ denotes the derivative of the interpolant at $x = x_i$. The diagonal matrices $\mathbb{A}_i$ and $\mathbb{B}_i$ in (23) are defined as follows

$$\mathbb{A}_i = diag\left\{ \alpha_i^1, \alpha_i^2, \dots, \alpha_i^m \right\}, \ \mathbb{B}_i = diag\left\{ \beta_i^1, \beta_i^2, \dots, \beta_i^m \right\}. \tag{24}$$

The cubic interpolant, $\mathbf{p} = \mathbf{p}(x)$, is component-wise monotone on $[x_i, x_{i+1}]$ *iff* one of the following conditions (e.g., [12], [28]) is satisfied:

$$\left(\alpha_i^k - 1 \right)^2 + \left(\alpha_i^k - 1 \right) \left(\beta_i^k - 1 \right) + \left(\beta_i^k - 1 \right)^2 - 3\left(\alpha_i^k + \beta_i^k - 2 \right) \leq 0, \tag{25}$$

$$\alpha_i^k + \beta_i^k \leq 3, \quad \alpha_i^k \geq 0, \ \beta_i^k \geq 0, \quad \forall i, k. \tag{26}$$

As reported in [28], the necessary and sufficient conditions for monotonicity of a C^1 piecewise cubic interpolant originally given by Ferguson and Miller (1969), and independently, by Fritsch and Carlson [12]. The region of monotonicity is shown in Figure 1b. The results of implementing a monotone C^1 piecewise cubic interpolation when compared with the classic cubic spline interpolation, are depicted in Figure 1a. We note (Figure 1a) that the constructed function produces monotone interpolation and this function coincides with the classic cubic spline at some sections where the classic cubic spline is monotone.
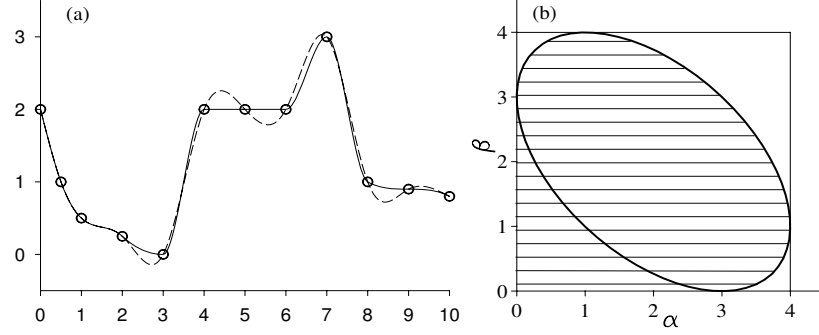
Figure 1: Monotone piecewise cubic interpolation. (a) Interpolation of a 1-D tabulated function. Circles: prescribed tabulated values; Dashed line: classic cubic spline; Solid line: monotone piecewise cubic. (b) Necessary and sufficient conditions for monotonicity. Horizontal hatching: region of monotonicity; Unshaded: cubic is non-monotone.

Using the cubic segment of the C^1 piecewise cubic interpolant, $\mathbf{p} = \mathbf{p}\,(x)$, (see, e.g., [12], [28]) for $x \in [x_i, x_{i+1}]$, we obtain the following interpolation formula

$$\mathbf{p}_{i+0.5} = 0.5\,(\mathbf{p}_i + \mathbf{p}_{i+1}) - \frac{\Delta x}{8}\left(\mathbf{p}'_{i+1} - \mathbf{p}'_i\right) + O\left((\Delta x)^r\right). \tag{27}$$

If $\mathbf{p}\,(x)$ has a continuous fourth derivative, then $r = 4$ in (27), see e.g. [27, p. 111]. However, the exact value of $\mathbf{p}'_i$ in (27) is, in general, unknown, and hence to construct numerical schemes, employing formulae similar to (27), the value of derivatives $\mathbf{p}'_i$ must be estimated.

Using (27) and the second formula in (22) we obtain from (21) the following scheme

$$\mathbf{v}_{i+0.5}^{n+0.5} = 0.5\left(\mathbf{v}_i^n + \mathbf{v}_{i+1}^n\right) - \frac{\Delta x}{8}\left(\mathbf{d}_{i+1}^n - \mathbf{d}_i^n\right) - \frac{\Delta t}{2}\frac{\mathbf{f}\left(\mathbf{v}_{i+1}^n\right) - \mathbf{f}\left(\mathbf{v}_i^n\right)}{\Delta x}, \tag{28}$$

where $\mathbf{d}_i^n$ denotes the derivative of the interpolant at $x = x_i$. In view of (27) and the second formula in (22), the local truncation error [32, p. 142], ψ, on a sufficiently smooth solution $\mathbf{u}(x,t)$ to (18) is found to be

$$\psi = O\left(\Delta t\right) + O\left(\frac{(\Delta x)^r}{\Delta t}\right) + O\left((\Delta t)^2 + (\Delta x)^2\right). \tag{29}$$

In view of (29) we conclude that the scheme (28) generates a conditional approximation, because it approximates (18) only if $(\Delta x)^r / \Delta t \to 0$ as $\Delta x \to 0$ and $\Delta t \to 0$. Let $\mathbf{d}_i^n$ be approximated with the accuracy $O\left((\Delta x)^s\right)$, then the value of r in (29) can be calculated (see Section 6, Proposition 5) by the following formula

$$r = \min\left(4, s + 1\right). \tag{30}$$

9

Interestingly, since (28) provides the conditional approximation, the order of accuracy depends on the pathway taken by Δx and Δt as $\Delta x \to 0$ and $\Delta t \to 0$. Actually, there exists a pathway such that Δt is proportional to $(\Delta x)^\mu$ and the CFL condition is fulfilled provided $\mu \geq 1$ and $\Delta x \leq \Delta x_0$, where Δx_0 is a positive value. If we take $\mu = 1$ and $s \geq 1$, then we obtain from (29) that the scheme (28) is of the first-order. If $\mu = 2$ and $s \geq 3$, then (28) is of the second-order. However, if $\mu = 2$ and $s = 2$, then, in view of (29) and (30), the scheme (28) is of the first-order. Moreover, under $\mu = 2$ and $s = 2$, the scheme will be of the first-order even if $\mathbf{g}_i^{n+0.25}$ in (21) will be approximated with the accuracy $O((\Delta t)^2)$. It seems likely that Example 6 in [30] can be seen as an illustration of the last assertion. The Nessyahu-Tadmor (NT) scheme with the second-order approximation of $\mathbf{d}_i^n$ is used [30] to solve a Burgers-type equation. Since $\Delta t = O((\Delta x)^2)$ [30], the NT scheme is of the first-order, and hence it can be the main reason for the scheme to exhibit the smeared discontinuity computed in [30, Fig. 6.22].

The approximation of derivatives $\mathbf{p}_i'$ can be done by the following three steps [12]: (i) an initialization of the derivatives $\mathbf{p}_i'$; (ii) the choice of subregion of monotonicity; (iii) modification of the initialized derivatives $\mathbf{p}_i'$ to produce a monotone interpolant.

The matter of initialization of the derivatives is the most subtle issue of this algorithm. Actually, the approximation of $\mathbf{p}_i'$ must, in general, be done with accuracy $O((\Delta x)^3)$ to obtain the second-order scheme when Δt is proportional to $(\Delta x)^2$, inasmuch as central schemes generate a conditional approximation. Thus, using the two-point or the three-point (centered) difference formula (e.g. [28], [41]) we obtain, in general, the first-order scheme. The so called limiter functions [28] lead, in general, to a low-order scheme as these limiters are often $O(\Delta x)$ or $O((\Delta x)^2)$ accurate. Performing the initialization of the derivatives $\mathbf{p}_i'$ in the interpolation formula (27) by the classic cubic spline interpolation [46], we obtain the approximation, which is $O((\Delta x)^3)$ accurate (e.g., [27], [28]), and hence, in general, the second-order scheme. The same accuracy, $O((\Delta x)^3)$, can be achieved by using the four-point approximation [28]. However, the efficiency of the algorithm based on the classic cubic spline interpolation is comparable with the one based on the four-point approximation, as the number of multiplications and divisions (as well as additions and subtractions) per one node is approximately the same for both algorithms. We will use the classic cubic spline interpolation for the initialization of the derivatives $\mathbf{P}_i'$ in the interpolation formula (27), as it is based on the tridiagonal algorithm, which is 'the rare case of an algorithm that, in practice, is more robust than theory says it should be' [46].

Obviously, for each interval $[x_i, x_{i+1}]$ in which the initialized derivatives $\mathbf{p}_i'$, $\mathbf{p}_{i+1}'$ such that at least one point (α_i^k, β_i^k) does not belong to the region of monotonicity (25)-(26), the derivatives $\mathbf{p}_i'$, $\mathbf{p}_{i+1}'$ must be modified to $\widetilde{\mathbf{p}}_i'$, $\widetilde{\mathbf{p}}_{i+1}'$ such that the point $(\widetilde{\alpha}_i^k, \widetilde{\beta}_i^k)$ will be in the region of monotonicity. The modification of the initialized derivatives, would be much simplified if we take a square

10

as a subregion of monotonicity. In connection with this, we will make use the subregions of monotonicity represented in the following form:

$$0 \leq \alpha_i^k \leq 4\aleph, \quad 0 \leq \beta_i^k \leq 4\aleph, \quad \forall i, k, \tag{31}$$

where $\aleph$ is a monotonicity parameter. Obviously, the condition (31) is sufficient for the monotonicity (see Figure 1b) provided that $0 \leq \aleph \leq 0.75$.

Let us now find necessary and sufficient conditions for (27) to be monotonicity preserving. By virtue of (23), the interpolation formula (27) can be rewritten to read

$$\mathbf{p}_{i+0.5} = \left(0.5\mathbf{I} + \frac{\mathbb{B}_i - \mathbb{A}_i}{8} \right) \cdot \mathbf{p}_i + \left(0.5\mathbf{I} - \frac{\mathbb{B}_i - \mathbb{A}_i}{8} \right) \cdot \mathbf{p}_{i+1}. \tag{32}$$

The coefficients of (32) will be non-negative *iff* $|\beta_i - \alpha_i| \leq 4$. Hence (27) will be monotonicity preserving *iff* (31) will be valid provided $0 \leq \aleph \leq 1$. Notice, there is no any contradiction between the sufficient conditions, (31) provided $0 \leq \aleph \leq 0.75$, for the interpolant, $\mathbf{p} = \mathbf{p}(x)$, to be monotone through the interval $[x_i, x_{i+1}]$, and the necessary and sufficient conditions, (31) provided $0 \leq \aleph \leq 1$, for the scheme (32) to be monotonicity preserving. In the latter case the interpolant, $\mathbf{p} = \mathbf{p}(x)$, may, in general, be non-monotone, however at the point $i + 0.5$ the value of an arbitrary component of $\mathbf{p}_{i+0.5}$ will be between the corresponding components of $\mathbf{p}_i$ and $\mathbf{p}_{i+1}$.

To fulfill the conditions of monotonicity (31), the modification of derivatives $\mathbf{p}_i' = \{p_i'^1, p_i'^2, \ldots, p_i'^m\}$ can be done by the following algorithm suggested, in fact, by Fritsch and Carlson [12] (see also [28]):

$$S_i^k := 4\aleph \min \mathrm{mod}(\Delta_{i-1}^k, \Delta_i^k), \quad \widetilde{p}_i'^k := \min \mathrm{mod}(p_i'^k, S_i^k), \quad \aleph = const, \tag{33}$$

where $\Delta_i^k = \left(p_{i+1}^k - p_i^k \right) / \Delta x$, the function $\min \mathrm{mod}(x, y)$ is defined (e.g., [28], [30], [36], [41], [50]) as follows

$$\min \mathrm{mod}(x, y) \equiv \frac{1}{2} \left[sgn(x) + sgn(y) \right] \min \left(|x|, |y| \right). \tag{34}$$

Let us note that instead of point values, $\mathbf{v}_{i+0.5}^n$, employed in the construction of the scheme (21), it can be used the cell averages (e.g., [5], [30], [32]) calculated on the basis of the monotone C^1 piecewise cubics. In such a case we obtain, instead of (27), the following interpolation formula

$$\mathbf{p}_{i+0.5} = 0.5 \left(\mathbf{p}_i + \mathbf{p}_{i+1} \right) - \varkappa \frac{\Delta x}{8} \left(\mathbf{p}_{i+1}' - \mathbf{p}_i' \right), \tag{35}$$

where $\varkappa = 2/3$. The region of monotonicity in this case will also be

$$0 \leq \mathbb{A}_i \leq 4\aleph\mathbf{I}, \ 0 \leq \mathbb{B}_i \leq 4\aleph\mathbf{I}, \quad 0 \leq \aleph \leq 1, \quad \forall i. \tag{36}$$

Notice, the interpolation formula (35) coincides with (27) under $\varkappa = 1$. Thus, in view of the interpolation formula (35), the staggered scheme (21) is written to read

$$\mathbf{v}_{i+0.5}^{n+0.5} = 0.5 \left(\mathbf{v}_{i+1}^n + \mathbf{v}_i^n \right) - \varkappa \frac{\Delta x}{8} \left(\mathbf{d}_{i+1}^n - \mathbf{d}_i^n \right) - \frac{\Delta t}{2} \frac{\mathbf{f} \left(\mathbf{v}_{i+1}^n \right) - \mathbf{f} \left(\mathbf{v}_i^n \right)}{\Delta x}, \tag{37}$$

11

where $\mathbf{d}_i^n$ denotes the derivative of the interpolant at $x = x_i$, the range of values for the parameter $\varkappa$ is the segment $0 \leq \varkappa \leq 1$. If $\varkappa = 1$ (or $\varkappa = 0$), then Scheme (37) coincides with the scheme (28) (or with the LxF scheme, respectively). As it was shown above, the scheme (37) is of the first order provided $\Delta t = O(\Delta x)$. The central scheme (37), approximating the 1-D equation (18) with the first order, will be abbreviated to as COS1.

4   Construction of central schemes

We will consider explicit schemes on a uniform grid with time step Δt and spatial mesh size Δx. In view of the CFL condition [32], we assume for the explicit schemes, that $\Delta t = O(\Delta x)$. Moreover, we will also assume that $\Delta x = O(\Delta t)$, since a central scheme generates a conditional approximation to Eq. (18) (see Section 3). In such a case, the following inequalities will be valid, for sufficiently small Δt and Δx,

$$\nu_0 \Delta t \leq \Delta x \leq \mu_0 \Delta t, \quad \nu_0, \mu_0 = const, \ 0 < \nu_0 \leq \mu_0. \tag{38}$$

Notice, for hyperbolic problems it is often assumed that Δt and Δx are related in a fixed manner (e.g., [32, p. 140], [47, p. 120]), i.e. it is assumed that Δt and Δx fulfill a more strong condition than (38).

Scheme (28) is of the first-order, $O(\Delta t + (\Delta x)^2)$, and non-oscillatory LxF scheme is of the first-order, $O(\Delta t + \Delta x)$. Let us demonstrate that (28) is, in fact, LxF scheme with a negative numerical viscosity added to obtain a higher order approximation to Eq. (18) with respect to x. We rewrite Scheme (37) to read

$$\frac{\mathbf{v}_{i+0.5}^{n+0.5} - \mathbf{v}_{i+0.5}^n}{0.5\Delta t} + \frac{\mathbf{f}\left(\mathbf{v}_{i+1}^n\right) - \mathbf{f}\left(\mathbf{v}_i^n\right)}{\Delta x} =$$
$$\frac{\Delta x^2}{\Delta t} \frac{\mathbf{v}_i^n - 2\mathbf{v}_{i+0.5}^n + \mathbf{v}_{i+1}^n}{\Delta x^2} - \varkappa \frac{\Delta x^2}{4\Delta t} \frac{\mathbf{d}_{i+1}^n - \mathbf{d}_i^n}{\Delta x}. \tag{39}$$

Notice, the second term in the right-hand side of (39) is, in fact, the negative numerical viscosity. Without this term ($\varkappa = 0$), Scheme (39) would be LxF scheme. As it is demonstrated in Section 5, Scheme (28) can exhibit spurious oscillations in contrast to LxF scheme. Interestingly, there is a possibility to improve Scheme (37) by introducing an additional positive numerical viscosity such that the scheme's order of accuracy would increase up to $O((\Delta t)^2 + (\Delta x)^2)$. Let us approximate $\mathbf{v}_{i+0.5}^n$ and $\mathbf{g}_i^{n+0.125}$ in (21) with the accuracy $O((\Delta x)^2 + (\Delta t)^2)$. Using Taylor series expansion, we write

$$\mathbf{g}_i^{n+0.25} = \mathbf{f}\left(\mathbf{v}_i^n\right) + \left.\frac{\partial \mathbf{f}\left(\mathbf{v}_i^n\right)}{\partial t}\right|_{t=t_n} \frac{\Delta t}{4} + O\left(\Delta t^2\right). \tag{40}$$

By virtue of the PDE system, (18), we find

$$\frac{\partial \mathbf{f}}{\partial t} = \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \cdot \frac{\partial \mathbf{u}}{\partial t} = -\frac{\partial \mathbf{f}}{\partial \mathbf{u}} \cdot \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \cdot \frac{\partial \mathbf{u}}{\partial x} = -\left(\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\right)^2 \cdot \frac{\partial \mathbf{u}}{\partial x}. \tag{41}$$

12

Using the interpolation formula (35) and the formulae (40)-(41), we obtain from (21) the following second order central scheme

$$\mathbf{v}_{i+0.5}^{n+0.5} = 0.5 \left(\mathbf{v}_{i+1}^n + \mathbf{v}_i^n\right) - \varkappa \frac{\Delta x}{8} \left(\mathbf{d}_{i+1}^n - \mathbf{d}_i^n\right) - \frac{\Delta t}{2} \frac{\mathbf{f}\left(\mathbf{v}_{i+1}^n\right) - \mathbf{f}\left(\mathbf{v}_i^n\right)}{\Delta x} +$$

$$\xi \frac{(\Delta t)^2}{8\Delta x} \left[\left(\mathbf{A}_{i+1}^n\right)^2 \cdot \mathbf{d}_{i+1}^n - \left(\mathbf{A}_i^n\right)^2 \cdot \mathbf{d}_i^n\right], \quad \mathbf{A} \equiv \frac{\partial \mathbf{f}}{\partial \mathbf{u}}, \tag{42}$$

where ξ is introduced by analogy with $\varkappa$ in (37), and hence $0 \leq \xi \leq 1$. Scheme (42) coincides with (37) provided that $\xi = 0$. Since $\mathbf{d}_i^n$ is the derivative of the interpolant at $x = x_i$, the last term in right-hand side of (42) can be seen as the non-negative numerical viscosity introduced into the first order scheme (37). Owing to this term, Scheme (42) is $O((\Delta x)^2 + (\Delta t)^2)$ accurate, provided that $\xi = 1$. Thus, we are dealing with the vanishing viscosity method [17], [32] and, hence, in view of [17, Theorem 3.3], the scheme, (42), satisfies the entropy condition. The central scheme (42), approximating the 1-D equation (18) with the second order, will be abbreviated to as COS2.

4.1 Stability of the second-order scheme COS2

In view of Theorem 4, the stability of (42) will be investigated on the basis of its variational scheme. It is assumed that the bounded operator $\mathbf{A}\ (= \partial \mathbf{f}(\mathbf{u})\,/\,\partial \mathbf{u})$ in (18) is Fréchet-differentiable on the set $\Omega_\mathbf{u} \subset \mathbb{R}^M$, and its derivative is bounded on $\Omega_\mathbf{u}$. Hence, the following inequalities are valid

$$\sup_{\mathbf{u} \in \Omega_\mathbf{u}} \|\mathbf{A}\| \leq \lambda_{\max} < \infty, \quad \|\delta \mathbf{A}_i^n\| = \left\|\frac{\partial \mathbf{A}_i^n}{\partial \mathbf{v}_i^n} \cdot \delta \mathbf{v}_i^n\right\| \leq \alpha_A \|\delta \mathbf{v}_i^n\|, \tag{43}$$

where $\lambda_{\max}$, $\alpha_A = const$. Considering that $\mathbf{v}_i^n$ in (42) is Lipschitz-continuous, we write

$$\left\|\mathbf{v}_i^n - \mathbf{v}_{i+1}^n\right\| \leq C_v \Delta x, \quad C_v = const. \tag{44}$$

By virtue of (23), the second term in right-hand side of (42) can be written in the form

$$\varkappa \frac{\Delta x}{8} \left(\mathbf{d}_{i+1}^n - \mathbf{d}_i^n\right) = \frac{\varkappa}{8} \left(\mathbb{B}_i^n - \mathbb{A}_i^n\right) \cdot \left(\mathbf{v}_{i+1}^n - \mathbf{v}_i^n\right). \tag{45}$$

Then, the variational scheme corresponding to (42) is the following

$$\delta \mathbf{v}_{i+0.5}^{n+0.5} = 0.5 \left(\delta \mathbf{v}_i^n + \delta \mathbf{v}_{i+1}^n\right) + \frac{\varkappa}{8} \left[\left(\mathbf{v}_i^n - \mathbf{v}_{i+1}^n\right)^T \cdot \delta \mathbb{D}_i^n\right]^T + \frac{\varkappa}{8} \mathbb{D}_i^n \cdot \left(\delta \mathbf{v}_i^n - \delta \mathbf{v}_{i+1}^n\right)$$

$$+\xi \frac{(\Delta t)^2}{8\Delta x^2} \left\{\left[\delta \left(\left(\mathbf{A}_{i+1}^n\right)^2 \cdot \mathbb{B}_i\right)\right] \cdot \left(\mathbf{v}_{i+1}^n - \mathbf{v}_i^n\right) - \left[\delta \left(\left(\mathbf{A}_i^n\right)^2 \cdot \mathbb{A}_i\right)\right] \cdot \left(\mathbf{v}_{i+1}^n - \mathbf{v}_i^n\right)\right\} +$$

$$\xi \frac{(\Delta t)^2}{8\Delta x^2} \left[\left(\mathbf{A}_{i+1}^n\right)^2 \cdot \mathbb{B}_i - \left(\mathbf{A}_i^n\right)^2 \cdot \mathbb{A}_i\right] \cdot \left(\delta \mathbf{v}_{i+1}^n - \delta \mathbf{v}_i^n\right) +$$

$$\frac{\Delta t}{2\Delta x} \left(\mathbf{A}_i^n \cdot \delta \mathbf{v}_i^n - \mathbf{A}_{i+1}^n \cdot \delta \mathbf{v}_{i+1}^n\right), \tag{46}$$

13

where $\mathbb{D}_i^n = diag\left\{D_{i,1}^n, D_{i,2}^n, \ldots, D_{i,M}^n\right\} \equiv \mathbb{B}_i^n - \mathbb{A}_i^n$. By virtue of (36), we find that $-4\aleph\mathbf{I} \leq \mathbb{D}_i^n \leq 4\aleph\mathbf{I}$, and hence $-8\aleph\mathbf{I} \leq \delta\mathbb{D}_i^n \leq 8\aleph\mathbf{I}$. Thus, we may write that

$$\|\delta\mathbb{D}_i^n\| \leq 8\aleph. \tag{47}$$

By virtue of (38), (47) and (44), and since $0 \leq \varkappa, \aleph \leq 1$, we find the following estimation for the second term in right-hand side of (46):

$$\left\|\frac{\varkappa}{8}\left[\left(\mathbf{v}_i^n - \mathbf{v}_{i+1}^n\right)^T \cdot \delta\mathbb{D}_i^n\right]^T\right\| \leq \frac{\varkappa}{8}\left\|\mathbf{v}_i^n - \mathbf{v}_{i+1}^n\right\| \|\delta\mathbb{D}_i^n\| \leq \mu_0 C_v \Delta t. \tag{48}$$

By virtue of (43), (38), (47), (44), and since $0 \leq \xi \leq 1$ and the CFL number $C_r = \Delta t \lambda_{\max}/\Delta x \leq 1$, we find the following estimation for the fourth and fifth terms in right-hand side of (46):

$$\left\|\xi\frac{(\Delta t)^2}{8\Delta x^2}\left[\delta\left(\left(\mathbf{A}_{i+1}^n\right)^2 \cdot \mathbb{B}_i\right)\right] \cdot \left(\mathbf{v}_{i+1}^n - \mathbf{v}_i^n\right)\right\| \leq \beta_A \Delta t \|\delta\mathbf{v}_i^n\|, \ \beta_A = const, \tag{49}$$

where β_A depends on the other constants, namely, on α_A, $\lambda_{\max}$, C_v, μ_0.

In view of (48) and (49), Scheme (46) will be stable if the following scheme be stable (see [48, pp. 390-392], [7, Theorem 7]).

$$\delta\mathbf{v}_{i+0.5}^{n+0.5} = 0.5\left(\delta\mathbf{v}_i^n + \delta\mathbf{v}_{i+1}^n\right) + \frac{\varkappa}{8}\mathbb{D}_i^n \cdot \left(\delta\mathbf{v}_i^n - \delta\mathbf{v}_{i+1}^n\right) +$$

$$\xi\frac{(\Delta t)^2}{8\Delta x^2}\left[\left(\mathbf{A}_{i+1}^n\right)^2 \cdot \mathbb{B}_i - \left(\mathbf{A}_i^n\right)^2 \cdot \mathbb{A}_i\right] \cdot \left(\delta\mathbf{v}_{i+1}^n - \delta\mathbf{v}_i^n\right) +$$

$$\frac{\Delta t}{2\Delta x}\left(\mathbf{A}_i^n \cdot \delta\mathbf{v}_i^n - \mathbf{A}_{i+1}^n \cdot \delta\mathbf{v}_{i+1}^n\right). \tag{50}$$

We rewrite (50) to read

$$\delta\mathbf{v}_{i+0.5}^{n+0.5} = 0.5\left(\mathbf{I} + \mathbf{E}_i^n\right) \cdot \delta\mathbf{v}_i^n + 0.5\left(\mathbf{I} - \mathbf{E}_{i+1}^n\right) \cdot \delta\mathbf{v}_{i+1}^n, \tag{51}$$

where

$$\mathbf{E}_i^n = \frac{\varkappa}{4}\mathbb{D}_i^n - \xi\frac{(\Delta t)^2}{4(\Delta x)^2}\left[\left(\mathbf{A}_{i+1}^n\right)^2 \cdot \mathbb{B}_i - \left(\mathbf{A}_i^n\right)^2 \cdot \mathbb{A}_i\right] + \frac{\Delta t}{\Delta x}\mathbf{A}_i^n, \tag{52}$$

$$\mathbf{E}_{i+1}^n = \frac{\varkappa}{4}\mathbb{D}_i^n - \xi\frac{(\Delta t)^2}{4(\Delta x)^2}\left[\left(\mathbf{A}_{i+1}^n\right)^2 \cdot \mathbb{B}_i - \left(\mathbf{A}_i^n\right)^2 \cdot \mathbb{A}_i\right] + \frac{\Delta t}{\Delta x}\mathbf{A}_{i+1}^n. \tag{53}$$

Since the operator $\mathbf{A}$ $(= \partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u})$ is Fréchet-differentiable, and its derivative is bounded, see (43), we get, by virtue of (44) and [40, Corollary 3.2.4], that

$$\left\|\mathbf{E}_{i+1}^n - \mathbf{E}_i^n\right\| = \frac{\Delta t}{\Delta x}\left\|\mathbf{A}_{i+1}^n - \mathbf{A}_i^n\right\| \leq \frac{\Delta t}{\Delta x}\alpha_A\|\delta\mathbf{v}_i^n\| \leq \alpha_A C_v \Delta t. \tag{54}$$

14

We find, in view of the first inequality in (43) and (36), that the spectrum $s\left(\mathbf{E}_i^n\right) \subset [-\lambda_E, \lambda_E]$, where

$$\lambda_E = \varkappa \aleph - \xi \left(\frac{\Delta t \lambda_{\max}}{\Delta x}\right)^2 \aleph + \frac{\Delta t}{\Delta x}\lambda_{\max}, \quad \forall i, n. \qquad (55)$$

Hence, by virtue of [7, Theorem 7] we find that the scheme (51) will be stable if

$$\max_{\lambda \in [-\lambda_E, \lambda_E]} 0.5\left(|1 + \lambda| + |1 - \lambda|\right) \leqslant 1, \quad \forall i, n. \qquad (56)$$

We obtain from (56) the following condition for the stability of the variational scheme (46)

$$\left(\varkappa - \xi C_r^2\right)\aleph + C_r \leq 1, \quad C_r = \frac{\Delta t \lambda_{\max}}{\Delta x} \leq 1. \qquad (57)$$

Thus, in view of Theorem 4 (see also [7, Theorem 3]), the scheme (42) will be stable if (57) be valid.

Let us note that the parameters $\varkappa$ and ξ are taken as constant in Scheme (42). However, in practice, it can be convenient to take that $\varkappa_i^n = \varkappa(\mathbf{v}_i^n)$ and $\xi_i^n = \xi(\mathbf{v}_i^n)$. In such a case the condition, (57), for the stability of (42) can be grounded in perfect analogy to the above, if

$$\|\delta \varkappa_i^n\| = \left\|\frac{\partial \varkappa_i^n}{\partial \mathbf{v}_i^n} \cdot \delta \mathbf{v}_i^n\right\| \leq \alpha_\varkappa \|\delta \mathbf{v}_i^n\|, \quad \|\delta \xi_i^n\| = \left\|\frac{\partial \xi_i^n}{\partial \mathbf{v}_i^n} \cdot \delta \mathbf{v}_i^n\right\| \leq \alpha_\xi \|\delta \mathbf{v}_i^n\|, \quad (58)$$

where $\alpha_\varkappa, \alpha_\xi = const.$

4.2   Operator splitting schemes

By virtue of the operator-splitting idea [6], [11], [18], [32] (see also LOS in [48]), the following chain of equations corresponds to the problem (1)

$$\frac{1}{2}\frac{\partial \mathbf{U}}{\partial t} = \frac{1}{\tau}\mathbf{q}\left(\mathbf{U}\right), \quad t_n < t \leq t_{n+0.5}, \quad \mathbf{U}\left(\mathbf{x}, t_n\right) = \mathbf{U}^n\left(\mathbf{x}\right), \qquad (59)$$

$$\frac{1}{2}\frac{\partial \mathbf{U}}{\partial t} + \sum_{j=1}^N \frac{\partial}{\partial x_j}\mathbf{f}_j\left(\mathbf{U}\right) = 0, \, t_{n+0.5} < t \leq t_{n+1}, \, \mathbf{U}\left(\mathbf{x}, t_{n+0.5}\right) = \mathbf{U}^{n+0.5}\left(\mathbf{x}\right), \quad (60)$$

where $\mathbf{U}^n\left(\mathbf{x}\right)$ denotes the solution to (60) at $t = t_n$, $\mathbf{U}^{n+0.5}\left(\mathbf{x}\right)$ denotes the solution to (59) at $t = t_{n+0.5}$. If a high-resolution method is used directly for the homogeneous conservation law (60), then it is natural to use a high-order scheme for (59). As applied to, in general, stiff ($\tau \ll 1$) System (1), the second order schemes can be constructed on the basis of operator-splitting techniques with ease if (59) will be approximated by an implicit scheme and (60) by an explicit one, see Proposition 6 in Section 6. As an example, let us develop a central scheme for a 1-D version of (1). After operator-splitting, the 1-D equation can be represented in the form

$$\frac{1}{2}\frac{\partial \mathbf{U}}{\partial t} = \frac{1}{\tau}\mathbf{q}\left(\mathbf{U}\right), \quad t_n < t \leq t_{n+0.25}, \quad \mathbf{U}\left(x, t_n\right) = \mathbf{U}^n\left(x\right), \qquad (61)$$

15

$$\frac{1}{2}\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial}{\partial x}\mathbf{f}\left(\mathbf{U}\right) = 0, \ t_{n+0.25} < t \leq t_{n+0.5}, \ \mathbf{U}\left(x, t_{n+0.25}\right) = \mathbf{U}^{n+0.25}\left(x\right). \ (62)$$

Let us first consider the case when the following first-order implicit scheme be used for (61)

$$\mathbf{v}_i^{n+0.25} = \mathbf{v}_i^n + \frac{\Delta t}{2\tau}\mathbf{q}\left(\mathbf{v}_i^{n+0.25}\right), \tag{63}$$

and a central scheme with nonstaggered grid cells will be used for (62). To eliminate the staggering in (37), we can define, e.g. [23], the nonstaggered cell-average as the average of its two neighboring staggered cell-averages. Then, by virtue of (37), we find

$$\mathbf{v}_i^{n+0.5} = 0.25\left(\mathbf{v}_{i-1}^{n+0.25} + 2\mathbf{v}_i^{n+0.25} + \mathbf{v}_{i+1}^{n+0.25}\right) - \varkappa\frac{\Delta x}{16}\left(\mathbf{d}_{i+1}^{n+0.25} - \mathbf{d}_{i-1}^{n+0.25}\right) -$$

$$\frac{\Delta t}{4\Delta x}\left(\mathbf{f}_{i+1}^{n+0.25} - \mathbf{f}_{i-1}^{n+0.25}\right). \tag{64}$$

It is clear that Scheme (64) approximates (62) with the accuracy $O(\Delta t + (\Delta x)^2)$, however, in view of Proposition 6 in Section 6, Scheme (63)-(64), taken as a whole, is of the second order approximation for the 1-D version of (1).

Let us develop another nonstaggered central scheme approximating a 1-D version of (1) with the accuracy $O((\Delta t)^2 + (\Delta x)^2)$ and such that its components (after operator splitting) will be of the second order. It can be done on the basis of the second order scheme (63), (64) with ease. Actually, adding to and subtracting from Equation (126) (see Section 6, Proposition 6), rewritten for $t_n < t \leq t_{n+0.5}$, the same quantity, we obtain (after operator splitting) the following scheme, instead of (63), (64),

$$\mathbf{v}_i^{n+0.25} = \mathbf{v}_i^n + \frac{\Delta t}{\tau}\mathbf{q}_i^{n+0.25} - \frac{(\Delta t)^2}{32}\left(\frac{\partial^2\mathbf{U}}{\partial t^2}\right)_i^{n+0.25}, \tag{65}$$

$$\mathbf{v}_i^{n+0.5} = 0.25\left(\mathbf{v}_{i-1}^{n+0.25} + 2\mathbf{v}_i^{n+0.25} + \mathbf{v}_{i+1}^{n+0.25}\right) - \varkappa\frac{\Delta x}{16}\left(\mathbf{d}_{i+1}^{n+0.25} - \mathbf{d}_{i-1}^{n+0.25}\right) +$$

$$\frac{(\Delta t)^2}{32}\left(\frac{\partial^2\mathbf{U}}{\partial t^2}\right)_i^{n+0.25} - \frac{\Delta t}{4\Delta x}\left(\mathbf{f}_{i+1}^{n+0.25} - \mathbf{f}_{i-1}^{n+0.25}\right). \tag{66}$$

Thus, Scheme (65) as well as Scheme (66) are of the second order, and Scheme (65)-(66), taken as a whole, is of the second order as well.

Using Taylor series expansion, and central differencing, we find

$$\mathbf{v}_i^{n+0.125} = \mathbf{v}_i^{n+0.25} - \frac{\Delta t}{8}\left(\frac{\partial\mathbf{U}}{\partial t}\right)_i^{n+0.25} +$$

$$\frac{1}{2}\left(\frac{\Delta t}{8}\right)^2\left(\frac{\partial^2\mathbf{U}}{\partial t^2}\right)_i^{n+0.25} + O\left((\Delta t)^3\right), \tag{67}$$

$$\mathbf{v}_i^{n+0.25} = \mathbf{v}_i^n + \frac{\Delta t}{4}\left(\frac{\partial\mathbf{U}}{\partial t}\right)_i^{n+0.125} + O\left((\Delta t)^3\right). \tag{68}$$

16

We obtain, by virtue of (41), (62), that

$$\frac{\partial^2 \mathbf{U}}{\partial t^2} = -2\frac{\partial}{\partial t}\left(\frac{\partial \mathbf{f}}{\partial x}\right) = -2\frac{\partial}{\partial x}\left(\frac{\partial \mathbf{f}}{\partial t}\right) = 4\frac{\partial}{\partial x}\left(\mathbf{A}^2 \cdot \frac{\partial \mathbf{U}}{\partial x}\right), \qquad (69)$$

where $\mathbf{A} = \partial \mathbf{f}/\partial \mathbf{U}$. Then

$$\left[\frac{\partial}{\partial x}\left(\mathbf{A}^2 \cdot \frac{\partial \mathbf{U}}{\partial x}\right)\right]_i^{n+0.25} = \frac{1}{\Delta x}\left[(\mathbf{A}_{i+0.5}^{n+0.25})^2 \cdot \frac{\mathbf{v}_{i+1}^{n+0.25} - \mathbf{v}_i^{n+0.25}}{\Delta x} - \right.$$

$$\left. (\mathbf{A}_{i-0.5}^{n+0.25})^2 \cdot \frac{\mathbf{v}_i^{n+0.25} - \mathbf{v}_{i-1}^{n+0.25}}{\Delta x}\right] + O\left((\Delta x)^2\right), \qquad (70)$$

where $(\mathbf{A}_{i+0.5}^{n+0.25})^2 = 0.5\left((\mathbf{A}_{i+1}^{n+0.25})^2 + (\mathbf{A}_i^{n+0.25})^2\right)$. By virtue of (59), (67)-(70), we rewrite Scheme (65)-(66) to read

$$\mathbf{v}_i^{n+0.125} = \mathbf{v}_i^{n+0.25} - \frac{\Delta t}{8\tau}\left(\mathbf{q}_i^{n+0.125} + \mathbf{q}_i^{n+0.25}\right), \qquad (71)$$

$$\mathbf{v}_i^{n+0.25} = \mathbf{v}_i^n + \frac{\Delta t}{2\tau}\mathbf{q}_i^{n+0.125}, \qquad (72)$$

$$\mathbf{v}_i^{n+0.5} = 0.25\left(\mathbf{v}_{i-1}^{n+0.25} + 2\mathbf{v}_i^{n+0.25} + \mathbf{v}_{i+1}^{n+0.25}\right) - \varkappa\frac{\Delta x}{16}\left(\mathbf{d}_{i+1}^{n+0.25} - \mathbf{d}_{i-1}^{n+0.25}\right) +$$

$$\frac{\xi(\Delta t)^2}{8(\Delta x)^2}\left[(\mathbf{A}_{i+0.5}^{n+0.25})^2 \cdot \left(\mathbf{v}_{i+1}^{n+0.25} - \mathbf{v}_i^{n+0.25}\right) - (\mathbf{A}_{i-0.5}^{n+0.25})^2 \cdot \left(\mathbf{v}_i^{n+0.25} - \mathbf{v}_{i-1}^{n+0.25}\right)\right]$$

$$-\frac{\Delta t}{4\Delta x}\left(\mathbf{f}_{i+1}^{n+0.25} - \mathbf{f}_{i-1}^{n+0.25}\right), \qquad (73)$$

where ξ is introduced in the third term in the right-hand side of (73) by analogy with Scheme (42), and, hence, $0 \leq \xi \leq 1$. If $\xi = 0$, then (73) coincides with (64), being $O(\Delta t + (\Delta x)^2)$ accurate. If $\varkappa = 1$ and $\xi = 1$, then Scheme (71)-(73) approximates the 1-D version of (1) with the accuracy $O((\Delta t)^2 + (\Delta x)^2)$.

By analogy with the scheme COS2, (42), we find the conditions for the stability of Scheme (73) using its scheme in variations. Scheme (73) will be stable if

$$\varkappa\aleph - 0.5\xi C_r^2 + C_r \leq 1, \quad C_r = \frac{\Delta t \lambda_{\max}}{\Delta x} \leq 1. \qquad (74)$$

Let us note that in practice (e.g., [32], [48]) the operator-splitting techniques find a wide range of application in designing economical schemes for Eq. (60) in the domains of complicated geometry. The resulting method, in general, will be only first-order accurate in time because of the splitting [32], [48]. Thus, in line with established practice we will replace the multidimensional Eq. (60) by the chain of the one-dimensional equations:

$$\frac{1}{2N}\frac{\partial \mathbf{U}_j}{\partial t} + \frac{\partial}{\partial x_j}\mathbf{f}_j(\mathbf{U}_j) = 0, \ t_{n+0.5+(j-1)/(2N)} < t \leq t_{n+0.5+j/(2N)}, \qquad (75)$$

17

where $\mathbf{U}_j\left(x, t_{n+0.5+(j-1)/(2N)}\right) = \mathbf{U}_{j-1}\left(x, t_{n+0.5+(j-1)/(2N)}\right)$, $j = 1, 2, \ldots N$, $\mathbf{U}_0\left(x, t_{n+0.5}\right)$ denotes the solution to (59) at $t = t_{n+0.5}$. Eq. (59) will be approximated by a first-order implicit scheme or a second-order implicit Runge-Kutta scheme. In particular, it will be used the following Runge-Kutta scheme

$$\mathbf{v}_i^{n+0.25} = \mathbf{v}_i^{n+0.5} - \frac{\Delta t}{2\tau}\mathbf{q}\left(\mathbf{v}_i^{n+0.5}\right), \ \mathbf{v}_i^{n+0.5} = \mathbf{v}_i^n + \frac{\Delta t}{\tau}\mathbf{q}\left(\mathbf{v}_i^{n+0.25}\right), \qquad (76)$$

since this scheme possesses a discrete analogy to the continuous asymptotic limit.

5 Exemplification and discussion

In this section, we are mainly concerned with verification of the second order central scheme COS2, (42).

5.1 Scalar non-linear equation

As the first stage in the verification, we will focus on the following scalar 1-D version of the problem (1):

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}f\left(u\right) = 0, \quad x \in \mathbb{R}, \ 0 < t \le T_{\max}; \quad u\left(x, t\right)|_{t=0} = u^0\left(x\right). \qquad (77)$$

We will solve the inviscid Burgers equation (i.e. $f\left(u\right) \equiv u^2/2$) with the following initial condition

$$u\left(x, 0\right) = \left\{ \begin{array}{ll} u_0, & x \in \left(h_L, h_R\right) \\ 0, & x \notin \left(h_L, h_R\right) \end{array} \right. , \quad h_R > h_L, \ u_0 = const \ne 0. \qquad (78)$$

The exact solution to (77), (78) is given by

$$u\left(x, t\right) = \left\{ \begin{array}{ll} u_1\left(x, t\right), & 0 < t \le T \\ u_2\left(x, t\right), & t > T \end{array} \right. , \qquad (79)$$

where $T = 2S/u_0$, $S = h_R - h_L$,

$$u_1\left(x, t\right) = \left\{ \begin{array}{ll} \frac{x-h_L}{b-h_L}u_0, & h_L < x \le b, \ b = u_0 t + h_L \\ u_0, & b < x \le 0.5u_0 t + h_R \\ 0, & x \le h_L \ or \ x > 0.5u_0 t + h_R \end{array} \right. , \qquad (80)$$

$$u_2\left(x, t\right) = \left\{ \begin{array}{ll} \frac{2S(x-h_L)}{(L-h_L)^2}u_0, & h_L < x \le L \\ 0, & x \le h_L \ or \ x > L \end{array} \right. , \qquad (81)$$

$$L = 2\sqrt{S^2 + 0.5u_0 S\left(t - T\right)} + h_L. \qquad (82)$$

First, it will be used Scheme (42) under $\xi = 0$, i.e. the first order in time central scheme COS1, (37). The numerical solutions were computed on a uniform grid with spatial increments of $\Delta x = 0.01$, the velocity $u_0 = 1$ in (78),
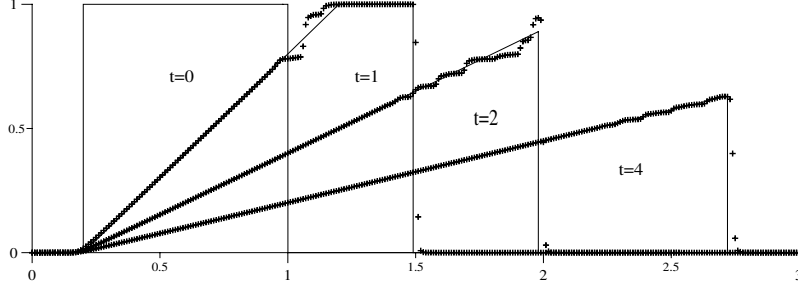
18

Figure 2: Inviscid Burgers equation. The scheme COS1 ($\varkappa = 1$) versus the analytical solution. Crosses: numerical solution; Solid line: analytical solution and initial data. $C_r = \aleph = 0.5$, $\Delta x = 0.01$.

$h_L = 0.2$, $h_R = 1$, the monotonicity parameter $\aleph = 0.5$, the CFL number $Cr \equiv u_0 \Delta t / \Delta x = 0.5$, and the parameter $\varkappa = 1$ in (37). The results of simulation are depicted with the exact solution in Figure 2. We note (Figure 2) that the first order scheme, (37), exhibits a typical second-order nature, however spurious solutions are produced by the scheme. Notice, the numerical simulations were performed with such values of the parameter $\varkappa$, CFL number, Cr, and monotonicity parameter, $\aleph$, that (57) was not violated. As it can be seen in Figure 2, the boundary maximum principle is not violated by the scheme, i.e., the maximum positive values of the dependent variable, v, occur at the boundary $t = 0$. It is interesting that the spurious solution (see Figure 2) produced by the scheme COS1 has the monotonicity property [19], since no new local extrema in x are created as well as the value of a local minimum is non-decreasing and the value of a local maximum is non-increasing.

Let us note that the problem of building free-of-spurious-oscillations schemes is, in general, unsettled up to the present. Even the best modern high-resolution schemes can produce spurious oscillations, and these oscillations are often of ENO type (see, e.g., [43] and references therein). We found that the oscillations produced by the COS1 scheme, (37), are of ENO type, namely their amplitude decreases rapidly with decreasing the time-increment Δt, and the oscillations virtually disappear under a relatively low CFL number, $Cr \leq 0.15$. However, the reduction of the CFL number causes some smearing of the solution. The spurious oscillations (see Figure 2) can be eradicated without reduction CFL number, but decreasing the parameter $\varkappa$. Particularly, the spurious oscillations disappear if $\varkappa = 2/3$, $C_r = 0.5$, however, this introduces more numerical smearing than in the case of the CFL number reduction. Satisfactory results are obtained under $\varkappa = 0.82$ ($C_r = 0.5$). The results of simulations are not depicted here.

To gain insight to why the scheme COS1, (37), can exhibit spurious solutions, let us consider the, so called, *first differential approximation* of this scheme ([13, p. 45], [49, p. 376]; see also 'modified equations' in [13, p. 45], [32], [35]). As reported in [13], [49], this heuristic method was originally presented by Hirt

19

(1968) (see [13, p. 45]) as well as by Shokin and Yanenko (1968) (see [49, p. 376]), and has since been widely employed in the development of stable difference schemes for PDEs.

We found that the local truncation error, ψ, for the scheme COS1 can be written in the following form

$$\psi = \frac{(1 - \varkappa)(\Delta x)^2}{4\Delta t} \frac{\partial^2 u(x,t)}{\partial x^2} + \frac{\Delta t}{4} \frac{\partial^2 f(u)}{\partial t \partial x} +$$

$$O\left(\frac{(\Delta x)^4}{\Delta t} + (\Delta t)^2 + (\Delta x)^2\right). \tag{83}$$

By virtue of (83), we find the first differential approximation of the scheme COS1

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \frac{\Delta t}{4} \frac{\partial}{\partial x}\left(B \frac{\partial u(x,t)}{\partial x}\right), \tag{84}$$

where $B = (1 - \varkappa)(\Delta x / \Delta t)^2 - A^2$. The term in right-hand side of (84) will be dissipative if

$$(1 - \varkappa)\left(\frac{\Delta x}{\Delta t}\right)^2 - A^2 > 0, \implies C_r^2 < 1 - \varkappa. \tag{85}$$

Thus, the scheme COS1, (37), is non-dissipative under $\varkappa = 1$, and hence can produce spurious oscillations. Notice, if $\varkappa = 0.82$, then we obtain from (85) that $C_r < 0.42$. Nevertheless, as it is reported above, satisfactory results can be obtained under $C_r = 0.5$ as well.

So then, the notion of first differential approximation has enabled us to understand that the spurious solutions exhibited by the scheme COS1, (37), are mainly associated with the negative numerical viscosity introduced to obtain the scheme of the second order in space, i.e. $O((\Delta x)^2 + \Delta t)$. Let us consider the scheme COS2, (42), approximating (77) with the accuracy $O((\Delta x)^2 + (\Delta t)^2)$. Notice, the second order scheme COS2, (42), is nothing more than the scheme COS1, (37), with the additional non-negative numerical viscosity. To test the scheme COS2, (42), the inviscid Burgers equation was solved under the initial condition (78). The numerical solutions were computed under the same values of parameters as in the case of the scheme COS1, but $C_r = 1$. The results of simulation are depicted with the exact solution in Figure 3.

We note (Figure 3) that the scheme COS2, (42), exhibits a typical second-order nature without any spurious oscillations. Increasing the value of $\aleph$ (up to $\aleph = 1$) leads to a minor improvement of the numerical solutions, whereas decreasing the value of Cr leads to a mild smearing of the solutions. The results of simulations are not depicted here.
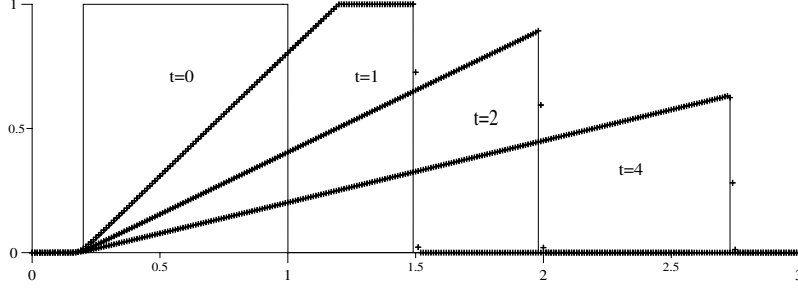
20

Figure 3: Inviscid Burgers equation. The scheme COS2 ($\xi = 1$, $\varkappa = 1$) versus the analytical solution. Crosses: numerical solution; Solid line: analytical solution and initial data. $C_r = 1$, $\aleph = 0.5$, $\Delta x = 0.01$.

5.2 Hyperbolic conservation laws with relaxation

Let us consider the model system of hyperbolic conservation laws with relaxation developed in [45]:

$$\frac{\partial w}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u^2 + aw\right) = 0, \tag{86}$$

$$\frac{\partial z}{\partial t} + \frac{\partial}{\partial x}az = \frac{1}{\tau}Q(w,z), \tag{87}$$

where

$$Q(w,z) = z - m(u - u_0), \quad u = w - q_0 z, \tag{88}$$

τ denotes the relaxation time of the system, q_0, m, a, and u_0 are constants. The Jacobian, $\mathbf{A}$, can be written in the form

$$\mathbf{A} = \left\{ \begin{array}{cc} w - q_0 z + a & -q_0\left(w - q_0 z\right) \\ 0 & a \end{array} \right\}. \tag{89}$$

The system (86)-(87) has the following frozen [45] characteristic speeds $\lambda_1 = a$, $\lambda_2 = u + a$. The equilibrium equation for (86)-(87) is

$$\frac{\partial w}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u_*^2 + aw\right) = 0, \tag{90}$$

where

$$u_* = w - q_0 z_*, \quad z_* = \frac{m}{1 + mq_0}\left(w - u_0\right). \tag{91}$$

The equilibrium characteristic speed λ_* can be written in the form

$$\lambda_*\left(w\right) = \frac{u_*\left(w\right)}{1 + mq_0} + a. \tag{92}$$

Pember's rarefaction test problem is to find the solution $\{w, z\}$ to (86)-(87), and hence the function $u = u\left(x, t\right)$, under $\tau \to 0$, and where

$$\{w, z\} = \left\{ \begin{array}{ll} \{w_L, z_*\left(w_L\right)\}, & x < x_0 \\ \{w_R, z_*\left(w_R\right)\}, & x > x_0 \end{array} \right., \tag{93}$$

21

$$0 < u_L = w_L - q_0 z_* \left(w_L \right) < u_R = w_R - q_0 z_* \left(w_R \right). \tag{94}$$

The analytical solution of this problem can be found in [45]. The parameters of the model system are assumed as follows: $q_0 = -1$, $m = -1$, $u_0 = 3$, $a = \pm 1$, $\tau = 10^{-8}$. The initial conditions of the rarefaction problem are defined by

$$u_L = 2, \implies z_L = m \left(u_L - u_0 \right) = 1, \ w_L = u_L + q_0 z_L = 1, \tag{95}$$

$$u_R = 3, \implies z_R = m \left(u_R - u_0 \right) = 0, \ w_R = u_R + q_0 z_R = 3. \tag{96}$$

The position of the initial discontinuity, x_0, is set according to the value of a so that the solutions of all the rarefaction problems are identical [45]. Let a position, x_R^t, of leading edge or a position, x_L^t, of trailing edge of the rarefaction be known (e.g., $x_R^t = 0.85$, $x_L^t = 0.7$ in [45]), then

$$x_0 = x_R^t - \left(\frac{u_R}{1 + mq_0} + a \right) t = x_L^t - \left(\frac{u_L}{1 + mq_0} + a \right) t. \tag{97}$$

At $t = 0.3$, under (95)-(96) we have [45]

$$u = \begin{cases} 2, & x \leq 0.7 \\ 2 + \frac{x - 0.7}{0.85 - 0.7}, & 0.7 < x < 0.85 \\ 3, & x \geq 0.85 \end{cases}. \tag{98}$$

The results of simulations, based upon the scheme COS2, (42), together with (76), under different values of the parameter a ($a = 1$, $a = -1$) and different values of a grid spacing, Δx, are depicted in Figure 4.

One can clearly see (Figure 4) that the scheme COS2 is free from spurious oscillations. Let us also note that the results generated by the scheme COS2 are less accurate in the case of negative value of a than those in the case of positive value of a. Specifically, in the numerical solutions produced under $a = -1$, the representations of the trailing and leading edges of the rarefaction are more smeared than those in the solutions produced under $a = 1$. Notice, under some negative value of a, the frozen and the equilibrium characteristic speeds do not all have the same sign.

5.3 1-D Euler equation of gas dynamics

In this subsection we apply the second order scheme COS2, (42), to the Euler equations of gamma-law gas:

$$\frac{\partial \mathbf{u} \left(x, t \right)}{\partial t} + \frac{\partial}{\partial x} \mathbf{F} \left(\mathbf{u} \right) = 0, \quad x \in \mathbb{R}, \ t > 0; \quad \mathbf{u} \left(x, 0 \right) = \mathbf{u}^0 \left(x \right), \tag{99}$$

$$\mathbf{u} \equiv \left\{ u_1, u_2, u_3 \right\}^T = \left\{ \rho, \rho v, e \right\}^T, \quad \mathbf{F} \left(\mathbf{u} \right) = \left\{ \rho v, \rho v^2 + p, \left(e + p \right) v \right\}^T, \tag{100}$$

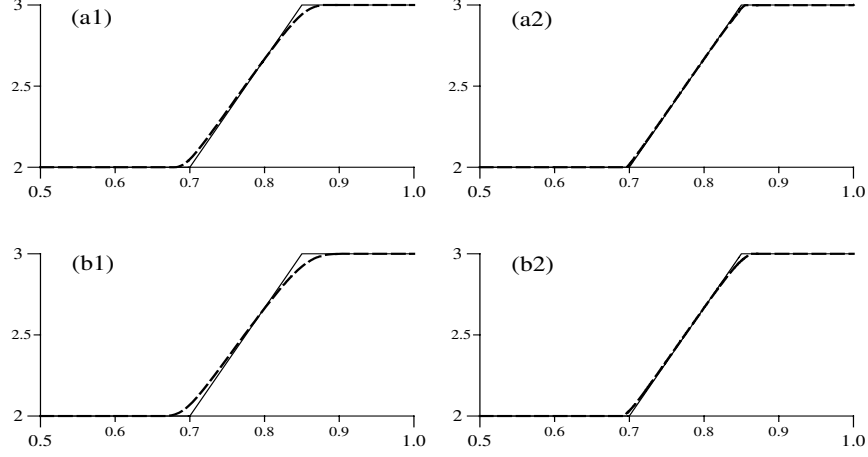$$e = \frac{p}{\gamma - 1} + \frac{1}{2} \rho v^2, \quad \gamma = const, \tag{101}$$

22

Figure 4: Pember's rarefaction test problem. The second-order scheme COS2 ($\xi = 1$, $\varkappa = 1$) versus the analytical solution for u. Dashed line: numerical solution; Solid line: analytical solution. Time $t = 0.3$, Courant number $C_r = 1$, monotonicity parameter $\aleph = 1$. (a1): $\Delta x = 10^{-3}$, $a = 1$; (a2): $\Delta x = 2.5 \times 10^{-4}$, $a = 1$; (b1): $\Delta x = 10^{-3}$, $a = -1$; (b2): $\Delta x = 2.5 \times 10^{-4}$, $a = -1$.

where ρ, v, p, e denote the density, velocity, pressure, and total energy respectively. We consider the Riemann problem subject to Riemann initial data

$$\mathbf{u}^0(x) = \left\{ \begin{array}{ll} \mathbf{u}_L & x < x_0 \\ \mathbf{u}_R & x > x_0 \end{array} \right. , \quad \mathbf{u}_L, \mathbf{u}_R = const. \tag{102}$$

The analytic solution to the Riemann problem can be found in [32, Sec. 14].

First we solve the shock tube problem (see, e.g., [5], [32], [33]) with Sod's initial data:

$$\mathbf{u}_L = \left\{ \begin{array}{c} 1 \\ 0 \\ 2.5 \end{array} \right\}, \quad \mathbf{u}_R = \left\{ \begin{array}{c} 0.125 \\ 0 \\ 0.25 \end{array} \right\}. \tag{103}$$

Following Balaguer and Conde [5] as well as Liu and Tadmor [33] we assume that the computational domain is $0 \leq x \leq 1$; the point x_0 is located at the middle of the interval $[0, 1]$, i.e. $x_0 = 0.5$; the equations (99) are integrated up to $t = 0.16$ on a spatial grid with 200 nodes as in [5] and in [33]. The CFL number is taken to be $Cr = 1$ in contrast to [5] and [33], where the simulations were done under $\Delta t = 0.1\Delta x$ (i.e. $0.13 \lesssim Cr \lesssim 0.22$). The results of simulations are depicted in Figure 5.

The results depicted in Figure 5 (left column) are not worse in comparison to the corresponding third-order central results of [33, p. 418] as well as to the results obtained by the fourth-order non-oscillatory scheme in [5, p. 472]. Notice, the fourth-order scheme [5, p. 472] gives a better resolution but, in contrast to the scheme COS2, can produce spurious oscillations.
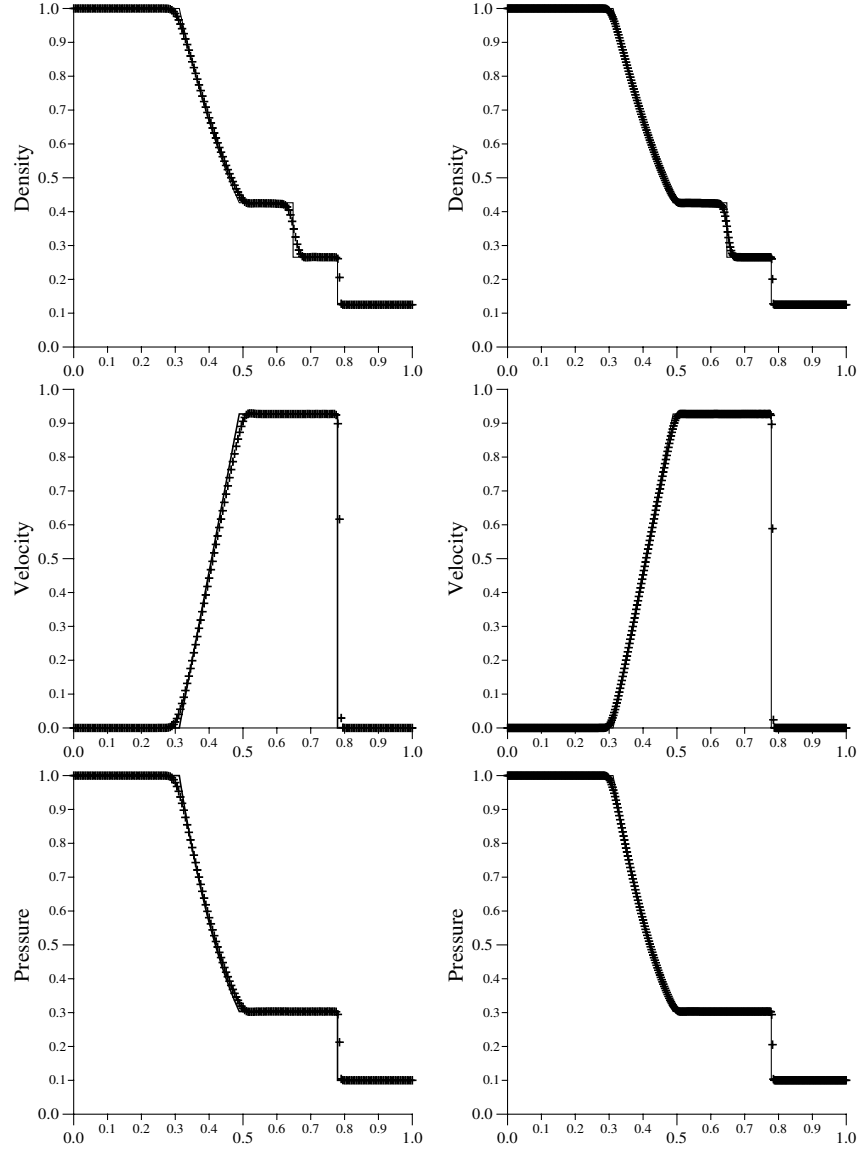
23

Figure 5: Sod's problem. The scheme COS2 under $C_r = 1$, $\aleph = 0.5$, $\xi = 1$, $\varkappa = 0.8$ versus the analytical solution. Time $t = 0.16$, spatial increment $\Delta x = 0.005$ (left column) and $\Delta x = 0.0025$ (right column).

24

Let us also note that the number of multiplications and divisions per one grid node in Scheme COS2, (42), and in the second-order scheme considered in [33] is approximately the same, but less than this number in the third- and fourth-order schemes developed in [33, p. 418], [5, p. 472], respectively. Hence, the results depicted in Figure 5 (right column) is rather chipper (in terms of CPU time) than the ones demonstrated in [33, p. 418], [5, p. 472]. Along with loss of computational efficiency, simulations with low CFL number can, in general, lead to excessive numerical smearing. As it is demonstrated above, Scheme COS2 is free from such drawbacks.

5.4   3-D axial symmetric gas dynamics

We consider an adiabatic expansion of a gas plume into vacuum [3], i.e., the so called Anisimov's problem. Taking into account the symmetry of the plume with respect to the axis z, the gas-dynamic equations can be written (for $0 < r$, $z < \infty$) as follows.

$$\frac{\partial \rho}{\partial t} + \frac{1}{r} \frac{\partial (r \rho v_r)}{\partial r} + \frac{\partial (\rho v_z)}{\partial z} = 0, \tag{104}$$

$$\frac{\partial}{\partial t} (\rho v_r) + \frac{1}{r} \frac{\partial}{\partial r} \left[r \rho (v_r)^2 \right] + \frac{\partial}{\partial z} (\rho v_z v_r) + \frac{\partial p}{\partial r} = 0, \tag{105}$$

$$\frac{\partial}{\partial t} (\rho v_z) + \frac{\partial}{\partial z} \left[\rho (v_z)^2 \right] + \frac{1}{r} \frac{\partial}{\partial r} (r \rho v_z v_r) + \frac{\partial p}{\partial z} = 0, \tag{106}$$

$$\frac{\partial \rho E}{\partial t} + \frac{1}{r} \frac{\partial}{\partial r} [r v_r (\rho E + p)] + \frac{\partial}{\partial z} [v_z (\rho E + p)] = 0. \tag{107}$$

$$\rho E = \frac{P}{\gamma - 1} + 0.5 \rho v^2, \ v^2 = v_r^2 + v_z^2, \ \gamma = const. \tag{108}$$

The initial conditions are the following (in details, see [3])

$$\rho = \rho (r, z), \ p / \rho^\gamma = const, \ v_r = v_z = 0, \ r, z \geq 0, \ t = 0. \tag{109}$$

At $r = 0$ we assume that the axis z is a reflection line. It prohibits any normal flux of mass through the boundary $r = 0$, i.e.

$$v_r = 0, \quad r = 0, \ z \geq 0. \tag{110}$$

Moreover, it is assumed that the pressure (p), density (ρ), and tangential velocity (v_z) are even functions of normal distance to the axis z while the normal velocity (v_r) is an odd function of r. It is also assumed that the plane $z = 0$ is a reflection surface, i.e. the pressure (p), density (ρ), and tangential velocity (v_r) are even functions of normal distance above the target surface while the normal velocity (v_z) is an odd function of z. The analytic solution to the problem (104)-(110) can be found in [3].

Notice, every point on the axis $r = 0$ is a singular point for System (104)-(107). Assuming that all terms at (104) are bounded values at a vicinity of $r = 0$, we find that $v_r \to 0$ as $r \to 0$. Hence

$$\lim_{r \to 0+0} \frac{\rho v_r|_{r>0}}{r} = \lim_{r \to 0+0} \frac{\rho v_r|_{r>0} - \rho v_r|_{r=0}}{r} = \frac{\partial (\rho v_r)}{\partial r}. \tag{111}$$

In perfect analogy we obtain

$$\frac{\rho\,(v_r)^2}{r} \to \frac{\partial\rho\,(v_r)^2}{\partial r},\ \frac{\rho v_z v_r}{r} \to \frac{\partial\rho v_z v_r}{\partial r},\ \frac{v_r\,(\rho E+p)}{r} \to \frac{\partial v_r\,(\rho E+p)}{\partial r}, \quad (112)$$

as $r \to 0$. By virtue of (111)-(112), we obtain from (104)-(107) the following conditions at $r = 0$:

$$\frac{\partial\rho}{\partial t} + \frac{\partial}{\partial r}\,(2\rho v_r) + \frac{\partial}{\partial z}\,(\rho v_z) = 0, \quad (113)$$

$$\frac{\partial}{\partial t}\,(\rho v_r) + \frac{\partial}{\partial r}\left[2\rho\,(v_r)^2 + p\right] + \frac{\partial}{\partial z}\,(\rho v_z v_r) = 0, \quad (114)$$

$$\frac{\partial}{\partial t}\,(\rho v_z) + \frac{\partial}{\partial z}\left[\rho\,(v_z)^2 + p\right] + \frac{\partial}{\partial r}\,(2\rho v_z v_r) = 0, \quad (115)$$

$$\frac{\partial\rho E}{\partial t} + \frac{\partial}{\partial r}\left[2v_r\,(\rho E+p)\right] + \frac{\partial}{\partial z}\left[v_z\,(\rho E+p)\right] = 0. \quad (116)$$

In the analytic solution [3] of the problem (104)-(110) the following input data are required: the initial dimensions of the plume, R_0 and Z_0, its mass M_P, and the initial energy E_P. We will use the following values as the reference quantities: $l_* = R_0$, $v_* = \sqrt{(5\gamma - 3)\,E_P/M_P}$, $t_* = l_*/v_*$, $\rho_* = M_P/\left(R_0^2 Z_0\right)$, $p_* = \rho_* v_*^2$.

The equations (104)-(107) are integrated up to $t = 0.4$ with $\sigma \equiv Z_0/R_0 = 0.1$. The CFL number is taken to be $Cr = 1$. It is assumed that the spatial increments are the following: $\Delta r = 0.0025$, $\Delta z = 0.00025$ if $0 < t \leq 0.05$; $\Delta r = 0.0025$, $\Delta z = 0.0005$ if $0.05 < t \leq 0.1$; $\Delta r = 0.005$, $\Delta z = 0.001$ if $0.1 < t \leq 0.2$; $\Delta r = 0.01$, $\Delta z = 0.002$ if $0.2 < t \leq 0.4$. The results of simulations as well as the analytical solution are depicted in Figures 6, 7. We observe (Figures 6, 7) that the numerical and analytical solutions are practically coincide, but in the vicinity of the front, namely, for very small values of density.

6 Appendix

Proposition 5 *Let us find the order of accuracy, r, in (27) if d_i will be approximated by $\widetilde{d}_i$ with the order of accuracy s, i.e. let*

$$d_i = \widetilde{d}_i + O\left((\Delta x)^s\right). \quad (117)$$

Let $U\,(x)$ be sufficiently smooth, then we can write

$$U_{i+1} = U_{i+05} + U'_{i+05}\frac{\Delta x}{2} + \frac{1}{2}U''_{i+05}\left(\frac{\Delta x}{2}\right)^2 + O\left((\Delta x)^3\right), \quad (118)$$

$$U_i = U_{i+05} - U'_{i+05}\frac{\Delta x}{2} + \frac{1}{2}U''_{i+05}\left(\frac{\Delta x}{2}\right)^2 + O\left((\Delta x)^3\right). \quad (119)$$
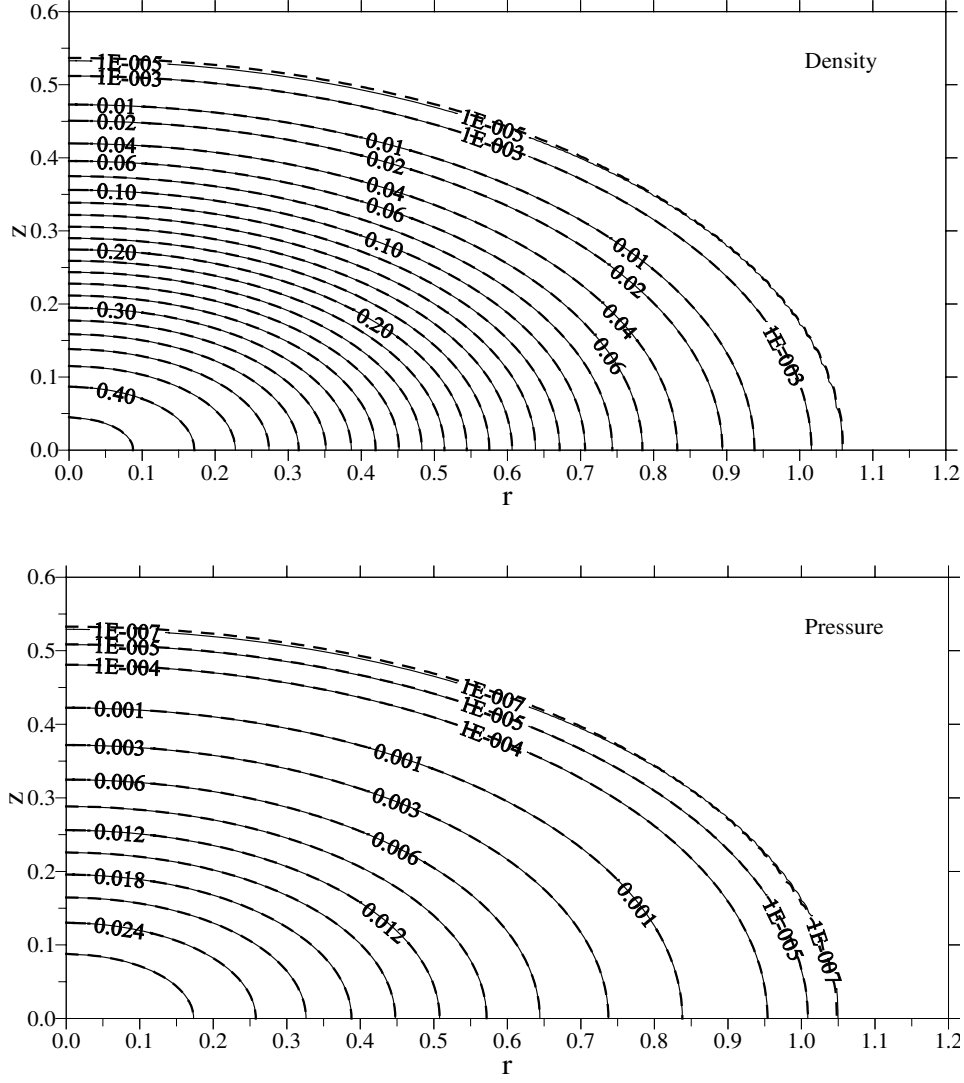
Figure 6: Anisimov's problem, density and pressure distribution. COS2 scheme versus the analytical solution. $\sigma \equiv Z_0 / R_0 = 0.1$, time $t = 0.4$, CFL number $C_r = 1$, monotonicity parameter $\aleph = \varkappa = 1$, spatial increments: $\Delta r = 0.0025$, $\Delta z = 0.00025$ if $0 < t \leq 0.05$; $\Delta r = 0.0025$, $\Delta z = 0.0005$ if $0.05 < t \leq 0.1$; $\Delta r = 0.005$, $\Delta z = 0.001$ if $0.1 < t \leq 0.2$; $\Delta r = 0.01$, $\Delta z = 0.002$ if $0.2 < t \leq 0.4$. Dashed lines: numerical solution; Solid lines: analytical solution.
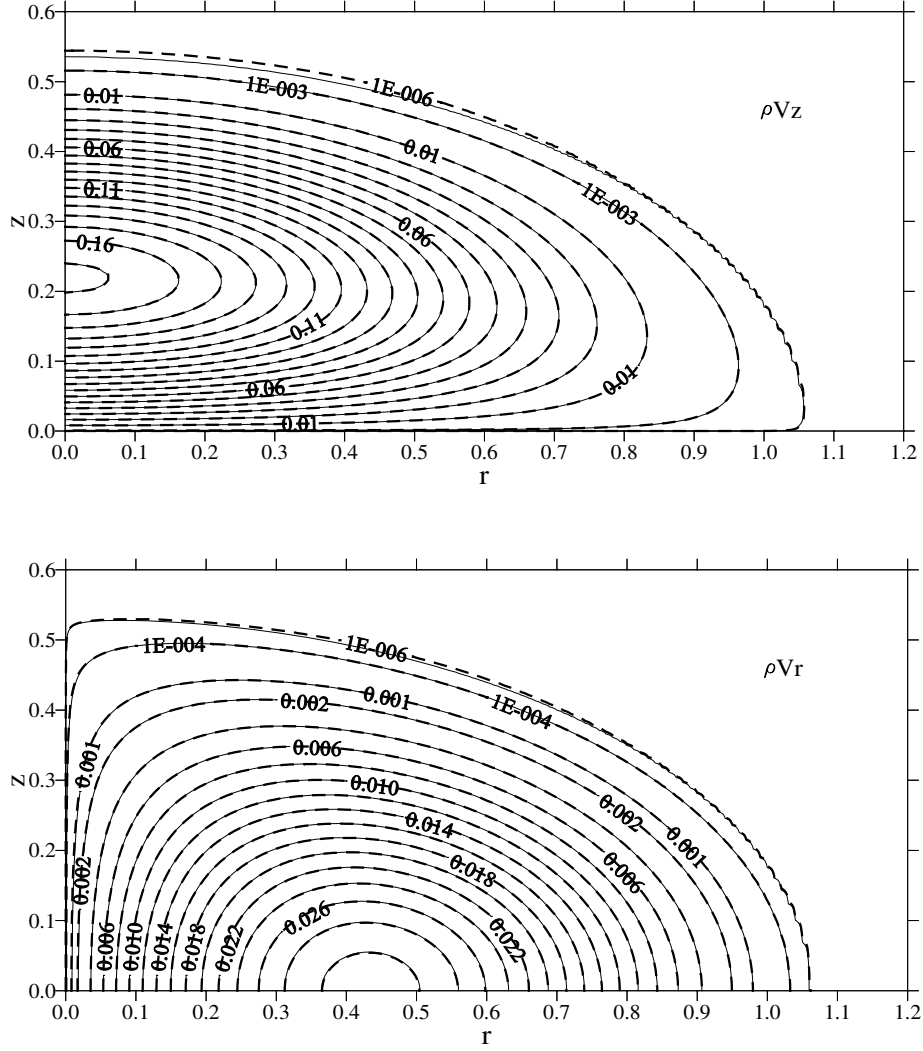
27

Figure 7: Anisimov's problem, momenta (ρV_z and ρV_r) distribution. COS2 scheme versus the analytical solution. $\sigma \equiv Z_0 / R_0 = 0.1$, time $t = 0.4$, CFL number $C_r = 1$, monotonicity parameter $\aleph = \varkappa = 1$, spatial increments: $\Delta r = 0.0025$, $\Delta z = 0.00025$ if $0 < t \leq 0.05$; $\Delta r = 0.0025$, $\Delta z = 0.0005$ if $0.05 < t \leq 0.1$; $\Delta r = 0.005$, $\Delta z = 0.001$ if $0.1 < t \leq 0.2$; $\Delta r = 0.01$, $\Delta z = 0.002$ if $0.2 < t \leq 0.4$. Dashed lines: numerical solution; Solid lines: analytical solution.

28

Combining the equalities (118) and 119 we obtain

$$U_{i+1} + U_i = 2U_{i+05} + \left.\frac{\partial^2 U}{\partial x^2}\right|_{i+05} \left(\frac{\Delta x}{2}\right)^2 + O\left((\Delta x)^3\right). \qquad (120)$$

In a similar manner we write:

$$d_{i+1} = U'_{i+05} + U''_{i+05}\frac{\Delta x}{2} + \frac{1}{2}U'''_{i+05}\left(\frac{\Delta x}{2}\right)^2 + O\left((\Delta x)^3\right), \qquad (121)$$

$$d_i = U'_{i+05} - U''_{i+05}\frac{\Delta x}{2} + \frac{1}{2}U'''_{i+05}\left(\frac{\Delta x}{2}\right)^2 + O\left((\Delta x)^3\right). \qquad (122)$$

Subtracting the equations (121) and (122), we obtain

$$\left.\frac{\partial^2 U}{\partial x^2}\right|_{i+05} = \frac{d_{i+1} - d_i}{\Delta x} + O\left((\Delta x)^2\right). \qquad (123)$$

In view of (123) and (117) we obtain from (120) the following interpolation formula

$$U_{i+05} = \frac{1}{2}\left(U_{i+1} + U_i\right) - \frac{\Delta x}{8}\left(\widetilde{d}_{i+1} - \widetilde{d}_i\right) + O\left((\Delta x)^4 + (\Delta x)^{s+1}\right). \qquad (124)$$

In view of (124) we obtain that $r = \min\left(4, s+1\right).$

Proposition 6 *Let us construct a second order scheme based on operator-splitting techniques. We will, in fact, use the summarized (summed) approximation method [48, Section 9.3] to estimate order of approximation. Consider the following equation*

$$\mathcal{P}\mathbf{u} \equiv \mathcal{P}_1\mathbf{u} + \mathcal{P}_2\mathbf{u} \equiv \frac{\partial \mathbf{u}}{\partial t} - L\mathbf{u} = 0, \quad \mathcal{P}_k\mathbf{u} \equiv \frac{1}{2}\frac{\partial \mathbf{u}}{\partial t} - L_k\mathbf{u}, \ k = 1, 2, \qquad (125)$$

where L_k is an operator, e.g. a differential operator, a real analytic function, etc., acting on $\mathbf{u}\left(x, t\right)$. We approximate (125) on the cell $[x_{i-1}, x_{i+1}] \times [t_n, t_{n+1}]$ by the following difference equation with the accuracy $O((\Delta x)^2 + (\Delta t)^2)$

$$\Pi\mathbf{v} \equiv \frac{\mathbf{v}_i^{n+1} - \mathbf{v}_i^n}{\Delta t} - \Lambda_1\mathbf{v}^{n+0.5} - \Lambda_2\mathbf{v}^{n+0.5} = 0, \qquad (126)$$

where it is assumed that the operator $L_k\mathbf{u}$ is approximated by the operator $\Lambda_k\mathbf{u}$ with the accuracy $O((\Delta x)^2)$, i.e.

$$\Lambda_k\mathbf{u}^{n+0.5} = (L_k\mathbf{u})_i^{n+0.5} + O\left((\Delta x)^2\right). \qquad (127)$$

In view of the operator splitting idea, to the problem (126) there corresponds the following chain of difference schemes

$$\Pi_1\mathbf{w} \equiv \frac{1}{2}\frac{\mathbf{w}_i^{n+0.5} - \mathbf{w}_i^n}{0.5\Delta t} - \Lambda_1\mathbf{w}_1^{n+0.5} = 0, \qquad (128)$$

$$\Pi_2 \mathbf{w} \equiv \frac{1}{2} \frac{\mathbf{w}_i^{n+1} - \mathbf{w}_i^{n+0.5}}{0.5\Delta t} - \Lambda_2 \mathbf{w}_2^{n+0.5} = 0. \tag{129}$$

One can see from the above that the operator $\mathcal{P}_k \mathbf{u}$ is approximated by $\Pi_k \mathbf{u}$ with the accuracy $O(\Delta t + (\Delta x)^2)$

$$\Pi_1 \mathbf{u}_i^{n+0.5} = (\mathcal{P}_1 \mathbf{u})_i^{n+0.5} - \frac{\Delta t}{8} \left(\frac{\partial^2 \mathbf{u}}{\partial t^2} \right)_i^{n+0.5} + O\left((\Delta t)^2 + (\Delta x)^2 \right), \tag{130}$$

$$\Pi_2 \mathbf{u}_i^{n+0.5} = (\mathcal{P}_2 \mathbf{u})_i^{n+0.5} + \frac{\Delta t}{8} \left(\frac{\partial^2 \mathbf{u}}{\partial t^2} \right)_i^{n+0.5} + O\left((\Delta t)^2 + (\Delta x)^2 \right). \tag{131}$$

In view of (130)-(131), the local truncation error [32, p. 142], ψ, on a sufficiently smooth solution $\mathbf{u}(x,t)$ to (125) is found to be

$$\psi = \Pi\mathbf{u} = \Pi_1\mathbf{u} + \Pi_2\mathbf{u} = \tag{132}$$

$$(\mathcal{P}_1\mathbf{u} + \mathcal{P}_2\mathbf{u})_i^{n+0.5} + O\left((\Delta t)^2 + (\Delta x)^2 \right) = O\left((\Delta t)^2 + (\Delta x)^2 \right). \tag{133}$$

Thus, the implicit scheme, (128), together with the explicit scheme, (129), approximate (125) with the second order.

References

[1] I. Ahmad., M. Berzins, MOL solvers for hyperbolic PDEs with source terms, Mathematics and Computers in Simulation 56 (2001) 115-125

[2] D. A. Anderson, J. C. Tannehill and R. H. Pletcher, *Computational Fluid Mechanics and Heat Transfer*, Hemisphere Publishing Corporation, New York, 1984.

[3] S. I. Anisimov, D. Bäuerle, B. S. Luk'yanchuk, Gas dynamics and film profiles in pulsed-laser deposition of materials, Phys. Rev. B 48 (16) (1993) 12 076-12 081.

[4] Mark A. Aves, David F. Griffiths, and Desmond J. Higham, Runge-Kutta solutions of a hyperbolic conservation law with source term, SIAM J. Sci. Comput., Vol. 22, No. 1, pp. 20-38 (2000)

[5] Angel Balaguer and Carlos Conde, Fourth-order non-oscillatory upwind and central schemes for hyperbolic conservation laws, SIAM J. Numer. Anal. , Vol. 43, No. 2, pp. 455-473, (2005)

[6] François Bereux, Lionel Sainsaulieu, A roe-type Riemann solver for hyperbolic systems with relaxation based on time-dependent wave decomposition, Numer. Math. 77: 143-185 (1997).

[7] V. S. Borisov and M Mond, On monotonicity, stability, and construction of central schemes for hyperbolic conservation laws with source terms, (2008), arXiv:0705.1109v4 [physics.comp-ph].

[8] V. S. Borisov and S. Sorek, - On monotonicity of difference schemes for computational physics, SIAM J. Sci. Comput., Vol. 25, No. 5 (2004), pp. 1557-1584.

[9] V. S. Borisov, On discrete maximum principles for linear equation systems and monotonicity of difference schemes, SIAM J. Matrix Anal. Appl., Vol. 24, No. 4 (2003), pp. 1110-1135.

[10] Russel E. Caflisch, Shi Jin, and Giovanni Russo, Uniformly accurate schemes for hyperbolic systems with relaxation, SIAM J. Numer. Anal., Vol. 34, No. 1 (1997), pp. 246-281.

[11] Tao Du, Jing Shi, Zi-Niu Wu, Mixed analytical/numerical method for flow equations with a source term, Computers & Fluids 32 (2003) 659-690.

[12] F.N. Fritsch and R.E. Carlson, Monotone piecewise cubic interpolation, SIAM J. Numer. Anal. 17, No. 2, 238-246 (1980).

[13] V. G. Ganzha and E. V. Vorozhtsov, *Numerical Solutions for Partial Differential Equations,* CRC Press, New York, 1996.

[14] V. G. Ganzha and E. V. Vorozhtsov, *Computer-aided analysis of difference schemes for partial differential equations*, John Wiley & Sons, New York, 1996.

[15] Gil' M. I., *Stability of finite and infinite dimensional systems*, Kluwer Academic Publishers, Boston, 1998.

[16] Gil' M. I., *Difference Equations in Normed Spaces, Stability and Oscillations*, Elsevier, Amsterdam, 2007.

[17] Edwige Godlewski and Pierre-Arnaud Raviart, *Numerical Approximation of Hyperbolic Systems of Conservation Laws,* Springer-Verlag, New York, 1996.

[18] L. Gosse, A Well-Balanced Flux-Vector Splitting Scheme Designed for Hyperbolic Systems of Conservation Laws with Source Terms, Computers and Mathematics with Applications 39 (2000) 135-159

[19] A. Harten, High resolution schemes for hyperbolic conservation laws, J. Comput. Phys., V. 49 (1983), pp. 357-393.

[20] Ami Harten, On a class of high resolution total-variaton-stable finite difference schemes, SIAM J. Numer. Anal., Vol. 21, No. 1 (1984), pp. 1-23.

[21] Ami Harten, Uniformly High Order Accurate Essentially Non-oscillatory Schemes, Ill, J. Comput. Phys., V. 71 (1987), pp. 231-303.

[22] Heinonen Juha, *Lectures on Lipschitz analysis*, Rep. Univ. Jyväskylä Dept. Math. Stat. 100 (2005), 1-77.

[23] G.-S. Jiang, D. Levy, C.-T. Lin, S. Osher, E. Tadmor, High-Resolution Nonoscillatory Central Schemes with Nonstaggered Grids for Hyperbolic Conservation Laws, SIAM Journal on Numerical Analysis, Vol. 35, No. 6. (Dec., 1998), pp. 2147-2168.

[24] Shi Jin, Runge-Kutta Methods for Hyperbolic Conservation Laws with Stiff Relaxation Terms, J. Comp. Phys. 122 (1995), 51-67.

[25] Shi Jin, Lorenzo Pareschi, and Gioseppe Toscani, Uniformly accurate diffusive schemes for multiscale transport equations, SIAM J. Numer. Anal. Vol. 38, No. 3 (2000), pp. 913-936.

[26] Shi Jin and C. David Levermore, Numerical Schemes for Hyperbolic Conservation Laws with Stiff Relaxation Terms, J. Comp. Phys. 126, 449-467 (1996).

[27] David Kahaner, Cleve Moler, and Stephen Nash, *Numerical methods and software*, Prentice-Hall, New Jersey, 1989.

[28] L.M. Kocić and G.V. Milovanović, Shape Preserving Approximations by Polynomials and Splines, Computers Math. Applic. Vol. 33, No. 11, pp. 59-97, 1997.

[29] Kolmogorov A. N., Fomin S. V. Introductory Real Analysis, Prentice Hall, USA, 1970.

[30] Alexander Kurganov and Eitan Tadmor, New High-Resolution Central Schemes for Nonlinear Conservation Laws and Convection-Diffusion Equations, Journal of Computational Physics 160, 241-282 (2000)

[31] Alexander Kurganov and Doron Levy, A third-order semidiscrete central scheme for conservation laws and convection-diffusion equations, SIAM J. Sci. Comput., Vol. 22, No. 4 (2000), pp. 1461-1488

[32] Randall J. LeVeque, *Finite volume methods for hyperbolic problems*, Cambridge University Press, Cambridge, 2002.

[33] Xu-Dong Liu and Eitan Tadmor, Third order nonoscillatory central scheme for hyperbolic conservation laws, Numer. Math. (1998) 79: 397-425.

[34] L. A. Monthé, A study of splitting scheme for hyperbolic conservation laws, Journal of Computational and Applied Mathematics 137 (2001) 1-12.

[35] K. W. Morton, *Numerical Solution of Convection-Diffusion Problems*, Chapman & Hall, London, 1996.

[36] K. W. Morton, Discretization of unsteady hyperbolic conservation laws, SIAM J. Numer. Anal., Vol. 39, No. 5 (2001), pp. 1556-1597.

[37] Giovanni Naldi and Lorenzo Pareschi, Numerical schemes for hyperbolic systems of conservation laws with stiff diffusive relaxation, SIAM J. Numer Anal., Vol. 37, No. 4 (2000), pp. 1246-1270.

[38] Greg F. Naterer and Jose A. Camberos, *Entropy Based Design and Analysis of Fluids Engineering Systems*,), Taylor and Francis Group, Boca Raton, USA, 2008.

[39] Haim Nessyahu and Eitan Tadmor, Non-oscillatory Central Differencing for Hyperbolic Conservation Laws, Journal of Computational Physics, Vol. 87, No 2., April 1990, pp. 408-463.

[40] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York, London, 1970.

[41] Lorenzo Pareschi, Central differencing based numerical schemes for hyperbolic conservation laws with relaxation terms, SIAM J. Numer. Anal., Vol. 39, No. 4 (2001), pp. 1395-1417.

[42] Lorenzo Pareschi and Giovanni Russo, Implicit-Explicit Runge-Kutta Schemes and Applications to Hyperbolic Systems with Relaxation, Journal of Scientific Computing, Vol. 25, Nos. 1/2, November 2005, pp. 129-155.

[43] Lorenzo Pareschi, Gabriella Puppo, and Giovanni Russo, Central Runge-Kutta Schemes for conservation laws, SIAM J. Sci. Comput., Vol. 26, No. 3 (2005), pp. 979-999.

[44] Richard B. Pember, Numerical methods for hyperbolic conservation laws with stiff relaxation I. Spurious solutions, SIAM J. Appl. Math., Vol. 53, No. 5, pp. 1293-1330, October 1993.

[45] Richard B. Pember, Numerical methods for hyperbolic conservation laws with stiff relaxation II. Higher-order Godunov methods, SlAM J. Sci. Comput., Vol. 14, No. 4, pp. 826-859, July 1993.

[46] William H. Press, Brian P. Flannery, Saul A. Teukolsky, William T. Vetterling, *Numerical Recipes in C*, The Art of Scientific Computing, Cambridge University Press, New York, 1988.

[47] R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems*, 2nd edn, Wiley-Interscience, New York, 1967.

[48] A. A. Samarskii, *The theory of difference schemes*, Marcel Dekker, New York, 2001.

[49] A. A. Samarskiy and A.V. Gulin, *Stability of Finite Difference Schemes*, Nauka, Moscow, 1973 (in Russian).

[50] Susana Serna and Antonio Marquina, Capturing shock waves in inelastic granular gases, Journal of Computational Physics 209 (2004) 787-795.